Identification and tracking of active RF-ID tags in an indoor environment

by

Thomas Michael McCoy

A thesis submitted to the University of London for the degree of Doctor of Philosophy

Faculty of Engineering
Department of Electronic & Electrical Engineering
University College London
The United Kingdom

October 2007

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Declaration of originality

I, Thomas Michael McCoy, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Name .... date .....4/2/08.....

ABSTRACT

Airport security has received widespread attention in recent years, becoming a critical issue in the eyes of the general public, security services and politicians alike. Over a similar period, the growing volume of flights, size of aircraft and the need, in major airports, to move a large number of people through a complex environment as efficiently and safely as possible is becoming an increasingly challenging task. With this in mind, the EU-funded Optag project was launched, in 2004, to study the feasibility of a combined RF-ID tag passenger location and panoramic video surveillance system to enhance airport efficiency, safety and security.

The basic concept is for airports to be fitted with a cellular network of combined RF-ID tag readers and high-resolution panoramic cameras, which are used to monitor the movements of people around the terminal building or buildings, locating each tag with some precision and helping to ensure that people are in the right place at the right time. This presents a number of technical and commercial challenges, which are addressed in this thesis by exploring a variety of approaches to the tag and reader design.

In this work, the design, development, and deployment of a prototype far field active RF-ID system is outlined. A tag is developed that incorporates a stable frequency source based on a carefully designed low phase noise synthesiser. One particular innovation explored by the author is a novel RF-ID receiver architecture incorporating frequency diversity that is able to accommodate low-cost tags with poor frequency stability. A custom communications protocol has been developed for the RF-ID system, which has been designed specifically to simultaneously accommodate and rapidly interrogate a large number of tags. A tag clashing analysis is presented and shows how pulse repetition interval diversity can be optimised to achieve maximum throughput of tags IDs.

The RF-ID reader design uses a differential RSSI AOA technique to locate the angular bearing of the tag and subsequent triangulation to establish location. The accuracy of this location technique has been assessed using a mixture of analytic, simulated and experimental techniques, culminating in a three cell trial based at UCL and Debrecen airport, Hungary. These trials have examined operation under a range of realistic conditions and have shown the system to be capable of a typical location accuracy of 1 m at 5 m range, and an operating range well in excess of the specified 20 m.

ACKNOWLEDGMENTS

This Ph.D research was funded by the European Union framework six.

I would like to express my gratitude to Professor P.V. Brennan for his help and supervision throughout this research. I would also like to thank Professor R. Benjamin for his advice during his visits to UCL. I am also grateful to Ben at Innovision R&T for the development of the enclosures for the readers and tags. I also appreciate the help from my colleagues Richard and Yanchuan in the radar group at UCL.

I would like to thank my family for their constant encouragement and support. Last, and far from least, I wish to thank Chin - I would never have finished this without you.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF PRINCIPAL SYMBOLS

Symbol	Definition	Units
A	Amplitude of carrier signal	V
B	Signal bandwidth	Hz
B_c	Coherent bandwidth	Hz
B_c	Compressor bandwidth	Hz
B_e	Expander bandwidth	Hz
c	Normalised received carrier power	V^2
$\frac{c}{n_o}$	Carrier power to noise power density ratio	dB Hz
d	Range	m
$\mathbf{E}$	Electric field vector	$V\ m^{-1}$
Δf_c	RX signal bandwidth	Hz
Δf_{if}	IF signal bandwidth	Hz
F	Receiver noise figure	dB
Δf_b	Detection bandwidth	Hz
f_c	Carrier frequency	Hz
f_{IF}	Final IF frequency	Hz
f_{res}	Frequency resolution	Hz
f_{ref}	Reference frequency	Hz
f_s	Sampling frequency	Hz
f_{sb}	Sampling bandwidth	Hz
G_r	Receiver gain	dB
G_t	Transmitter gain	dB
k	Boltzmanns constant	JK^{-1}
L	Free space path loss	none
SNR	Signal to noise power ratio	none

N	Number of FFT samples	none
N	Total division ratio	none
NF	Noise figure	dB
N_{opt}	Optimum repetition rate	s^{-1}
N_{cycles}	Symbol cycles in FFT	none
n	No. of far-field RFID tags in reader cell	none
P	No. of Pre-amble symbols	none
P	Received carrier power	V^2
P_r	received power	dBm
P_t	transmited power	dBm
R	Data rate	Hz
t	Signal propagation time	s
T	Temperature	K
T_b	Symbol time	s
T_d	Delay spread	s
T_c	Compressor time length	s
T_e	Expander time length	s
T_{mean}	Mean update time	s
T_{opt}	Optimum repetition interval	s
T_r	Pulse repetition interval	s
T_n	Pulse repetition rate	s^{-1}
T_s	Sampling time	s
T_w	Pre-amble time length	s
V_n	Voltage scaling constant	V
V_1 and V_2	Adjacent channel voltages	V
w_c	Carrier centre frequency	rads
w_n	Loop centre frequency	rads s^{-1}
w_{ref}	Reference frequency	rads s^{-1}
W	communication bandwidth	Hz
X	Channel unbalance	ratio

ε_r	Dielectric constant	none
ε	Complex permittivity	F/m
ϕ_{od}	phase deviation	rads
ϕ	Phase margin	rads
ϕ	Offset angle from boresight	(°)
ϕ_o	Phase constant	rads
$\Delta\phi$	Peak phase deviation	rads
Ψ	Angular displacement from cross-over point	(°)
λ	Wavelength	m
μ	Permeability	H/m
μ	Chirp rate	Hz s^{-1}
η_o	Noise spectral density	V^2 Hz^{-1}
σ_{coarse}	Coarse FFT frequency error	Hz
σ^2	Rms noise power in a FFT bin	V^2
τ_n	Time delay	s
τ	Time duration of data burst	s
ψ	Ray phase shift	rads
ψ_o	Initial ray phase shift	rads
$\Delta\Psi$	Angular bearing error	(°)
$\hat{\omega}_c$	Carrier frequency estimate	rads

ABBREVIATIONS

RFID	Radio frequency identification
AOA	Angle of arrival
RSS	Received signal strength
TOA	Time of arrival
TDOA	Time difference of arrival
LOS	Line of sight
NLOS	Non Direct line of sight
PLF	Polarisation loss factor
AF	Antenna factor
DF	Direction finding
dBc	Power in decibels relative to the carrier
dBm	Power in decibels relative to 1mW
dBpp	Power in decibels relative to the peak to peak
VCO	Voltage controlled oscillator
AGC	Analog gain controller
DF	Direction finding
DRO	Dielectric resonator oscillator
FET	Field effect transistor
PCB	Printed circuit board
MUSIC	MUltiple SIgnal Classification
SNR	Signal to noise power ratio
BPE	Best point estimate
RSSI	Received signal strength indicator
MDS	Minimum discernible signal
FSK	Frequency shift keyed
PSK	Phase shift keyed
FSL	Free space loss
LNA	Low noise amplifier
IRM	Image reject mixer

1. INTRODUCTION

1.1 Area of study

The subject of this thesis is the identification and tracking of people in an indoor environment using far field radio frequency identification (RF-ID) technology. Radio-frequency identification (RF-ID) is a term coined for short-range radio technology used to communicate mainly digital information between a stationary location and a movable object or between movable objects. These systems have become common in places where access control and tracking of physical objects is required. Examples include cattle herding, car immobilisers, and, most notably transport ticketing [1]. More recently, RF-ID systems have begun to find greater use in the consumer object identification market, in industrial automation, and in supply chain management. The use of RF-ID systems in these application domains has been promoted by efforts to develop low cost RF-ID tags as an economical replacement of bar-codes [2]. RF-ID is also now being actively considered as a method for locating people in an indoor environment, for instance at airports [3].

The fundamental purpose of airports is the efficient movement of passengers and freight. At present, the growing volume of flights, size of aircraft and the need, in major airports, to move a large number of people through a complex environment as efficiently and safely as possible is becoming an increasingly challenging task. The enormity of this task has led to a desire to improve boarding efficiency in order to reduce the occurence of flight delays. Over a similar period, airport security has received increasing attention with a strong desire to introduce effective security measures that do not unduly inhibit normal airport operations. With this in mind, the EU-funded Optag project was launched, in 2004, to study the feasibility of a combined RF-ID tag passenger location and video surveillance system to enhance airport efficiency, safety and security.

The project which is entitled 'Optag: Improving airport efficiency, security and passenger flow by enhanced passenger monitoring' aims to improve airport efficiency and security through the active tracking of passengers and staff using Radio Frequency Identification (RF-ID) tags combined with advanced high definition digital CCTV monitoring. The system aims to facilitate real-time location of individual passengers within the airport, the analysis of both mass traffic & individual behaviours and, where appropriate, the semi-automatic control of CCTV based vision systems to observe and record suspicious or unauthorised activity. The proposed system also enables the location of checked in passengers who are either missing or late, and thus reduce passenger-induced delays. Passenger or baggage problems have been identified by a recent feasibility study as causing around 5% of aircraft departure delays and costing approximately €150 million p.a. in Europe [3].

1.2 The 'OPTag' concept

OPTag is an RF-ID tag based video tracking and location system [4]. The system, as illustrated in Figure 1.1, comprises a cellular network of base-stations each consisting of a cluster of video cameras providing 360° azimuthal or hemispherical coverage that return a complete panoramic image or sub-section of such image to a central station. The base stations also receive and identify the ID of a number of RF tags within a small radius and can establish the bearing of each tag. The video images, tag IDs, respective bearings and signal strengths are returned via data links to a central control and gathering station at which the information is stored and processed. The system works in conjunction with a set of RF tags which may be numerous, perhaps totalling a thousand or more within each cell, each with a unique ID which is linked to a person, or item. A given area is covered by a network of base-stations which is capable of locating the position of each and every tag and providing video from the appropriate camera to image and track the vicinity of a given tag or tags.

The function of the system is to locate and record the position of tags within the area defined by the base-stations with the ability to automatically achieve video tracking and surveillance of selected tags. Typically, the cells may be of around

10 m radius. Direction finding is incorporated within each base station to further define the position of each tag and allow the appropriate camera to provide an image oriented towards the desired tag. The location of the tags is updated regularly on a real-time plan of the terminal using a local user interface, such as envisaged in Figure 1.2. The system software may also be written in such a way that it can be expanded to perform a wide range of tasks, drawing on the entire tag database and their associated locations. For example in a Public Address function, in which the operator may communicate with people in a given cell or cells to give directed announcements, for instance, to notify late-running passengers of imminent flight departures. The development of a highly polished and versatile system and associated software is envisaged for future commercialisation, which would encompass multiple operators performing a variety of disparate functions ranging from security to crowd control.

An airport terminal building application is illustrated in Figure 1.3. The RF tag here is very compact, low cost and mass-produced to the extent of being a disposable item. It may be embedded in, for instance, in boarding cards or luggage tags or incorporated in a wrist band and issued routinely to every passenger and all personnel. The RF tag emits a unique ID, one of four billion, with suitable coding, which is received and de-coded, at regular intervals of say 15-30 seconds, at the nearby base stations. The tag may utilise spread spectrum and/or pulsed techniques in order to allow hundreds of tags to be uniquely identified within a single cell without problems of interference. The base-station also contains a means for RF direction finding, which may comprise amplitude and/or phase comparison from a set of antennas mounted on the camera unit. The camera unit comprises multiple cameras capable of forming a panoramic image, perhaps over a complete hemisphere.

For each tag detected, the base station returns the ID, signal strength and bearing, along with the base-station ID, to a central control and data-gathering station at which the data is stored and processed. The operators of the station may be responsible for airport security, safety and customer support services in which they

may use the system to monitor and/or image a given tagged individual simply by entering his or her name or tag ID, thus allowing the automatic video tracking of the tagged individual around the complete area; monitor and co-ordinate crowd control; locate late-running passengers and make targeted announcements to them; and identify, locate, video image and communicate with any lost person. Data collected from all tags in the terminal building may be stored in computer memory for retrospective analysis as required.

1.2.1 Potential Applications

Other applications of the system are numerous, and may include:

(1) Lost person location, in shopping centres, airports, theme parks and so on, in which the location of a lost person may be established and video imaging automatically provided to assist in their identification and to, for instance, determine whether they are in distress.

(2) Crowd control and monitoring, in for example, shopping centres, sports stadiums, theme parks and airports, in which the flow of people may be recorded and statistically analysed, to assist in future planning and safety considerations.

(3) Market research in shopping centres, theme parks, exhibition centres to establish the popularity of given retail outlets, stands or attractions, and the average duration of stay.

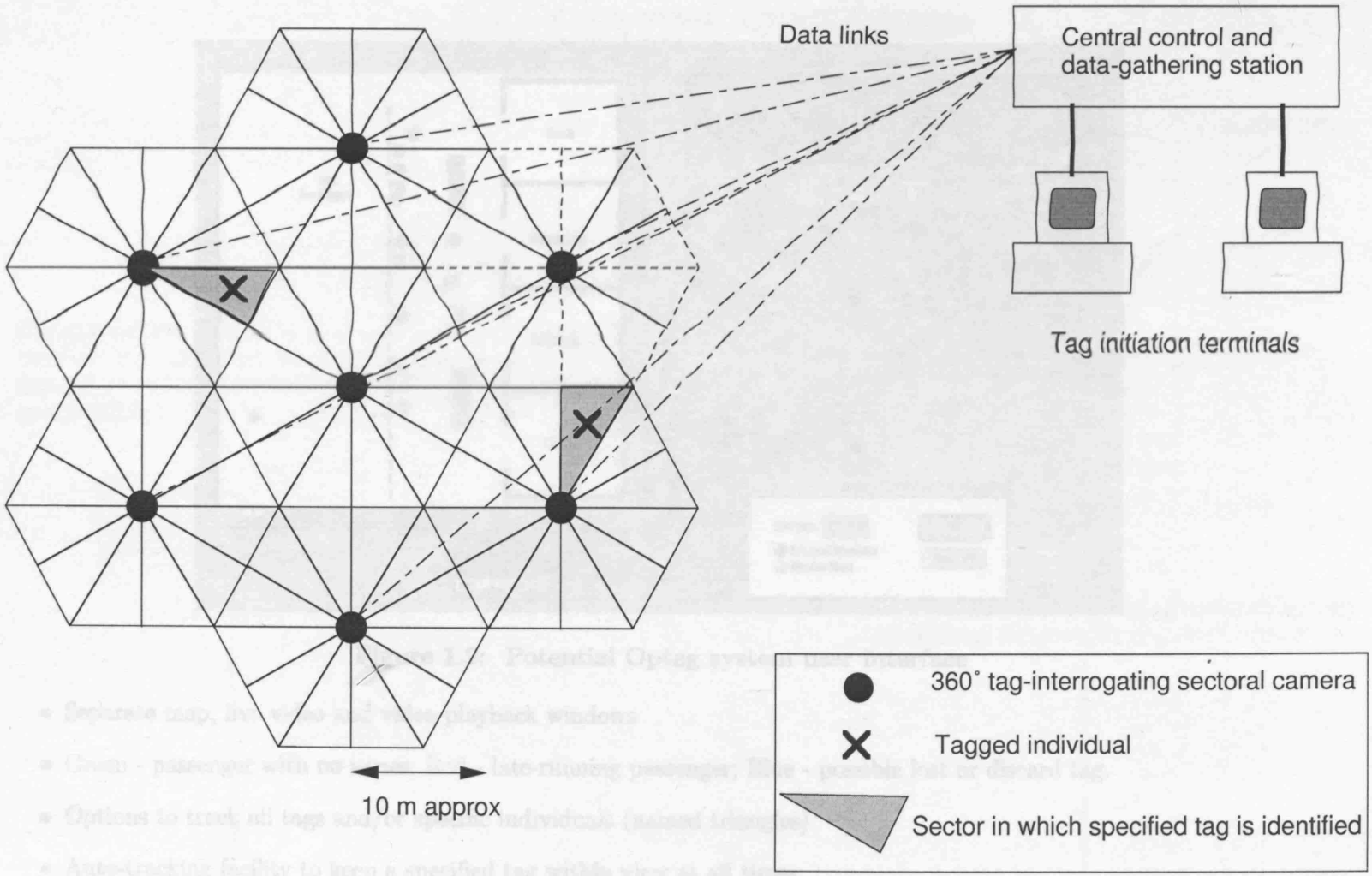


Figure 1.1: Proposed layout of RF-ID cell

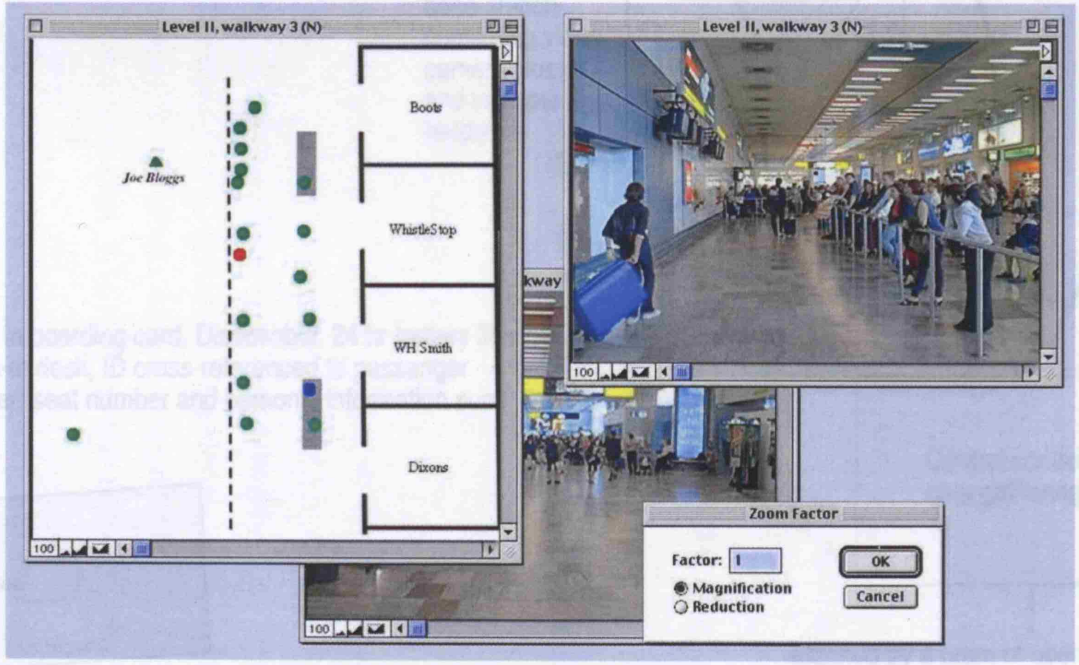


Figure 1.2: Potential Optag system user interface

- Separate map, live video and video playback windows
- Green - passenger with no issues; Red - late-running passenger; Blue - possible lost or discard tag.
- Options to track all tags and/or specific individuals (named triangles)
- Auto-tracking facility to keep a specified tag within view at all times

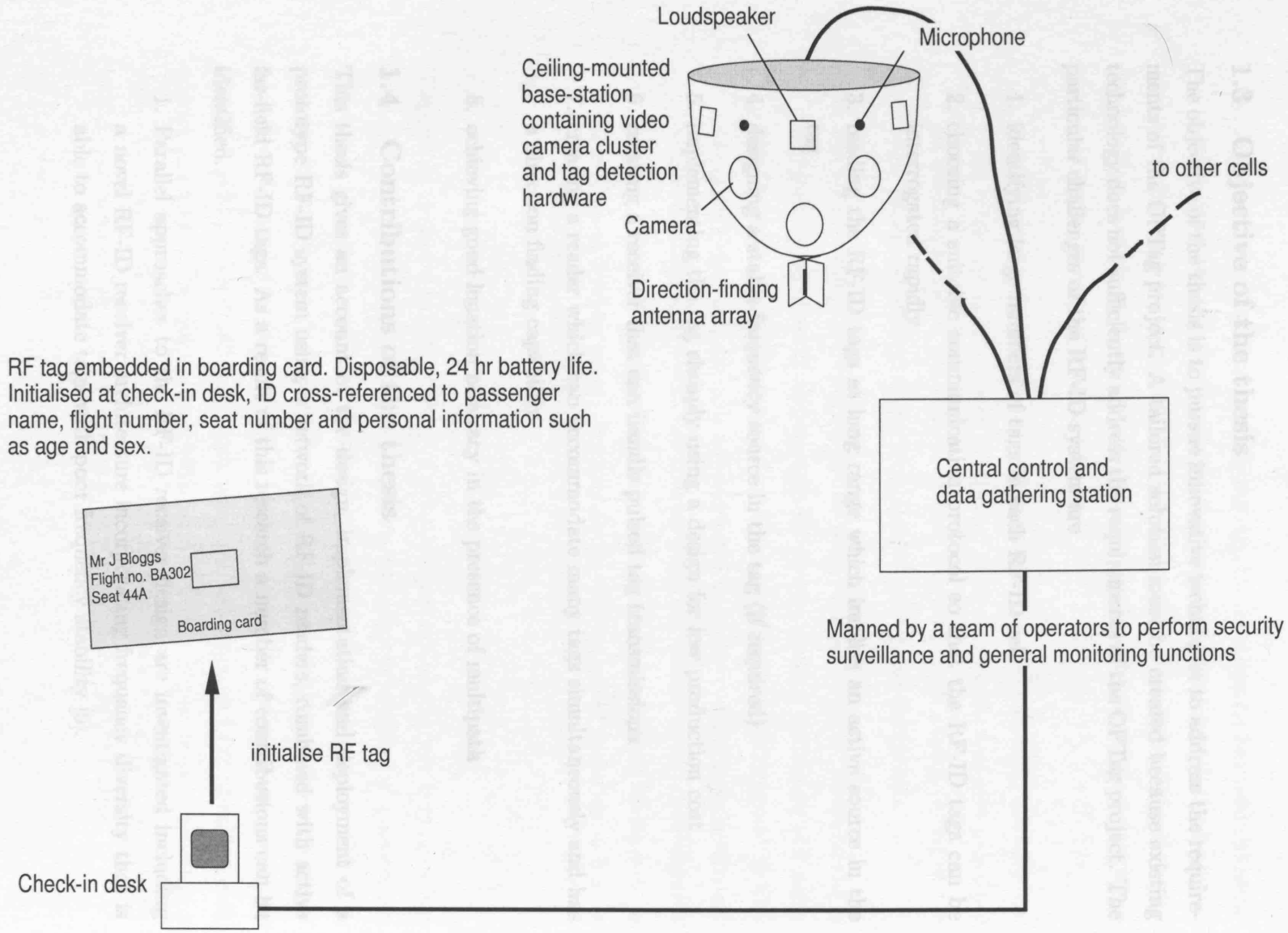


Figure 1.3: An example application - airport terminal building security and surveillance

1.3 Objective of the thesis

The objective of the thesis is to pursue innovative techniques to address the requirements of the OPTag project. A tailored solution must be created because existing technology does not sufficiently address the requirements of the OPTag project. The particular challenges of the RF-ID system are

1. identifying large numbers of tags in each RF-ID cell
2. choosing a suitable communication protocol so that the RF-ID tags can be interrogated rapidly
3. reading the RF-ID tags at long range which implies an active source in the tag
4. designing a stable frequency source in the tag (if required)
5. implementing the tag cheaply using a design for low production cost
6. creating a receiver that can handle pulsed tag transmissions
7. creating a reader which can accommodate many tags simultaneously and has a direction finding capability
8. achieving good location accuracy in the presence of multipath

1.4 Contributions of the thesis

This thesis gives an account of the design, implementation and deployment of a prototype RF-ID system using a network of RF-ID readers, combined with active far-field RF-ID tags. As a result of this research a number of contributions can be identified.

1. Parallel approaches to the RF-ID receiver design are investigated including a novel RF-ID receiver architecture incorporating frequency diversity that is able to accommodate tags with poor frequency stability [5].

2. Tag clashing analysis is developed showing the relationship between the mean update time and the the length of the tag transmission interval, the length of the tag burst and the number of tags in the cell. The analysis is then extended to show how the pulse repetition interval diversity can be optimised to achieve maximum throughput of tags IDs.

3. A synthesised stable tag frequency source has been designed and built with low phase noise and an optimised synthesiser design. Analysis was also presented on the adaptation of the PLL to provide frequency shift keying (FSK) of the VCO tuning input so that the loop dynamics were not disturbed.

4. Direction finding performance of the passive antenna array arrangement is analysed taking account of propagation effects, channel unbalance and additive Gaussian noise to provide a realistic assessment of performance in this demanding environment and experimental validation. Significantly, the experiments have shown that the corner-mounted reader with a two element receiver array performs much better than a central-mounted reader with a four element receiver array. Angular errors with standard deviation as low as 16 degrees are possible when the corner mounting is chosen.

1.5 Thesis layout

The overall structure takes the form of seven chapters. Chapter two begins with a brief examination of prior attempts at tagging and airport tracking of passsengers followed by an overview of RF-ID technology. The chapter then goes on to discuss location finding in an indoor environment. This involves assessing the advantages and disadvantages of existing direction finding techniques. In addition a number of experimental systems using RF technology are examined and the merits of each system analysed. The chapter concludes with a review of ray tracing techniques used in the numerical simulation of RF channel propagation that may be of value in the evaluation of the location finding performance of the Optag system.

Chapter three considers the implementation of the active RF-ID system. A link budget analysis is first performed to determine the likely operating range of the system. Next the cell capacity is evaluated using the clashing analysis of the simplex communication protocol adopted by the RF-ID system. The chapter then provides an extensive investigation of a novel RF-ID receiver architecture incorporating frequency diversity. The theoretical merits of this wide band synchroniser are investigated along with the hardware implementation including the design of the carrier frequency estimator using SAW (surface acoustic wave) dispersive delay lines.

Chapter four shows the development of the RF-ID system to be used in the Optag trial. The chapter comprises three sections. It begins with the design and development of the RF-ID tag and is then followed by the design and development of the rather more complex RF-ID receiver. The chapter finishes by examining the performance of the passive direction finding antenna array. The design and development of the antenna elements used in the array is also described here.

Chapter five describes the application of propagation modelling to location estimation. The first half of the chapter introduces two models representing the RF-ID system operating in an indoor environment. One is a simple 2D model and the other a more complex 3D model incorporating additional features not present in the 2D model. The remainder of the chapter describes how the reader topology and the averaging algorithm used on the angular bearings affect the location error.

Chapter six presents experimental results of the deployment of the RF-ID system in two locations. The performance of the RF-ID system is assessed using a series of tests. As well as measuring the mean and rms angular error for the tag wearer positioned on a regular grid when there were no obstructions, the RF-ID system also examines variations in location accuracy whenever there is obstruction and crowding.

The final chapter draws upon the entire thesis, tying up the various theoretical and experimental strands in order to present a summary of the work performed and the possibilities for the future.

2. BACKGROUND

2.1 Current State of the Art

The Optag system is unqiue in that two complementary yet independent systems (one RFID based, one vision based) are combined to ensure an appropriate balance of active and passive tracking, to enable an operator to track a passenger throughout the facility and to maximise security of the system. There is no existing working system which gives the benefits of the Optag system. However, systems have been demonstrated which show some of the benefits. A number of airports have experimented with techniques to accelerate passenger check-in or tracking through an airport. To date these have been based on short range (passive) tags, or longer-range active tags of various forms. These are summarised as follows.

2.1.1 Tagging and Airport Passenger Tracking

A trial was held at Zurich airport [1] in 1999 using short range readable passive tags (or smart cards) where a passenger has to pass through a gate or pinch point in order for it to be read. The benefits in terms of improved passenger flow were such that the original pilot trial in Zurich Airport apparently led to plans to adopt the system in Paris and Brussels airports, though this did not materialise. This system did not provide any information on passenger location so it would not assist with tracing lost passengers or identifying suspicious security behaviours. The maximum range of a typical passive tag would be around 1 metre or less and so the gates and pinch points must be sited very carefully to ensure accurate reliable reading of the information in the tag.

Longer-range active tags based on Bluetooth or similar radio technologies have been attempted. Perhaps the best known of the Bluetooth passenger trials was

[1] Zurich Airport contactless e-pass cards allow passengers to register their flight and identity details by passing through the gates. The information is confirmed at passport control where another gate is situated.

carried out at Billund airport by the BlueTags company [2]. In this trial, a Bluetooth transceiver is carried by a passenger who has pre-registered on the airport database giving personal information, flight information and a mobile phone number. On arrival at the airport, the tag is seen and the passenger can be identified and tracked with the relevant information being displayed to passport, security or gate personnel as required. The tag information can be used to offer passengers special promotions in the airport shops or restaurants, which are transmitted to them by SMS on their mobile phone. It is believed that the disadvantages of this system chiefly stem from the high bulk and cost of the passenger worn transceiver, and the lack of visual checks. This is a voluntary system which gives a frequent flyer passenger benefits in terms of speed of processing, but is not applicable to every passenger in an airport.

An airline has previously performed a pilot trial of passenger tracking based on their GSM mobile phone. [3] This relies on the airport IT system having access to the mobile phone operator's database and the consent of the passenger. The passenger's approach to the airport can be spotted, and their presence in a particular GSM cell can be reported. The intention was that micro GSM cells would be located within terminals so that a loosely defined zone can be identified. The purpose of this trial was more to predict flight take up and improve the flight filling statistics than to track passengers. In the current regulatory framework, tracking by mobile phone signals requires that the customer opts-in to such a service and so it is unlikely to afford coverage of complete passenger populations in the medium term.

2.1.2 RF-ID Tags

RF-ID (Radio Frequency Identification) is a technique for providing a unique means of identifying (tagging) an object or person by non-contact electronic means. The object of any RF-ID system is to carry data in suitable transponders, generally known as tags, and to retrieve data, by machine-readable means, at a suitable time and place to satisfy particular application needs. Data within a tag may provide identification of an item in manufacture, goods in transit, a location, or the identity

[2]BlueTags A/S Denmark

[3]Conversation of Bob Lloyd, OPTag project co-ordinator, with member of BA staff.

of a vehicle, animal or individual.

Communication of data between tags and a reader is by wireless means. Two methods distinguish and categorise RF-ID systems, one based upon close proximity reactive coupling and one based upon propagating electromagnetic waves. Coupling is via 'antenna' structures forming an integral feature in both tags and readers. While the term antenna is generally considered more appropriate for propagating systems it is also loosely applied to inductive systems. The speed and accuracy of transmitting data is influenced by the variability and influences of the media or channels through which the data has to pass, including the air interface. There are a number of applications of RF-ID tagging systems and these operate at different frequencies which have different characteristics. Table 2.1 illustrates briefly some of the key aspects of RF-ID technology presently available. More detailed information about the standards and frequency allocation for short range devices can be found at http://www.etsi.org and http://www.ero.dk [1].

Table 2.1: Comparison of existing RF-ID Tagging technology

Frequency Band	Characteristics (Current technology)	Typical Applications
Low 100-500 kHz	Short to medium read range Inexpensive Low reading speed	Access control Animal identification Inventory control Car immobiliser
Intermediate 10-15 MHz	Short to medium read range Inexpensive Medium reading speed	Access control Smart cards Most widely used
High 850-950 MHz 2.4-5.8 GHz	Long read range High reading speed Line of sight required Expensive	Railroad car monitoring Toll collection systems

2.2 Indoor position finding

The accurate identification and indoor location of people and objects has become the focus of renewed interest since the proliferation of wireless technologies. The most widely used electronic navigation system is the global positioning system (GPS) which uses a network of satellites to provide a very accurate estimate of the location of a receiver on earth. However, in an indoor environment GPS works very poorly due to the attenuation and distortion by multipath propagation of the microwave signals at the receiver. This has led to a need for an effective indoor RF wireless location system since the precise tracking of an RF transducer in these circumstances is a difficult problem for all location systems and has no entirely satisfactory solution at present. This is because of errors due to propagation effects such as reflection, refraction, scattering and possible obstruction of line of sight before it reaches a set of disparate scattered receivers. This complex engineering problem has been approached by many academic communities including networking, robotics, vision and signal processing.

2.2.1 A brief review of location techniques

Figure 2.1 illustrates the typical elements of a location system, comprising physical sensors, positioning algorithm and display driver.

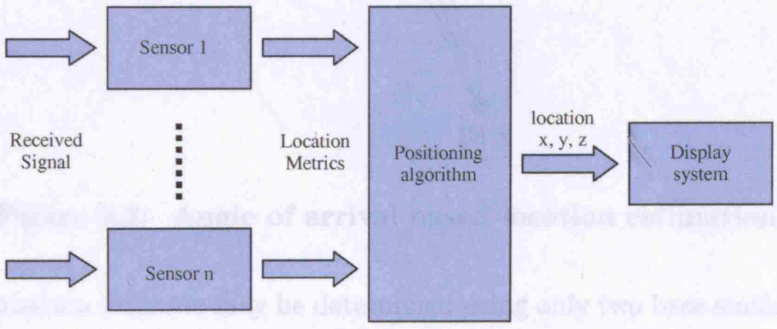


Figure 2.1: Functional block diagram of a wireless geolocation system

An introduction to the location techniques is now presented. There are three main aproaches angle of arrival (AOA), received signal strength (RSS) and time based measurements – time of arrival (TOA) and time difference of arrival (TDOA).

Beginning with AOA, as shown in Figure 2.2 the location of the transmitter can be found from the overlapping region of the bearings from two or more base stations located at known points, so-called triangulation. If there is an uncertainty of $\pm\phi_s$ in the true bearing angle from each base station then the base station will restrict the transmitter to a position inside the beam around the dashed line of sight (LOS) signal path with an angular spread of $2\phi_s$. The true transmitter location may then be found anywhere inside a polygon, known as a 'cocked' hat [6]. The estimation of the AOA, commonly referred to as direction finding (DF), can be accomplished with an array of antennas with narrow beamwidths. Amplitude comparison and phase comparison are the two methods commonly used to calculate the AOA.

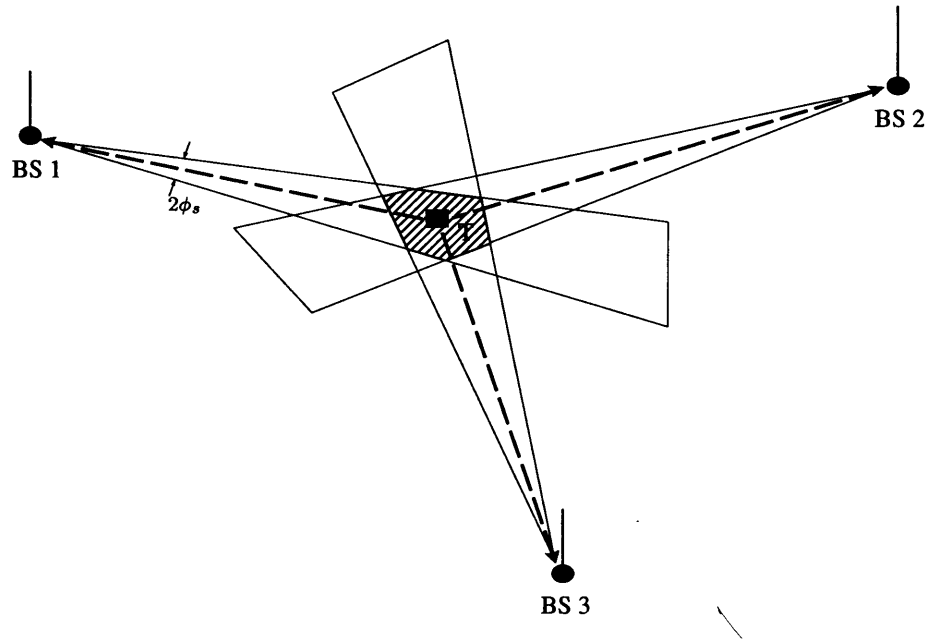


Figure 2.2: Angle of arrival based location estimation

The position estimate may be determined using only two base stations as long as the transmitter is not on the co-planar line joining the two base stations. There is also no requirement for time synchronisation between base stations, which is another major benefit of using AOA techniques. One disadvantage of AOA is that the position estimate degrades as function of the transmitter distance and position with respect to the base station [7]. RSS and TOA methods rely on a lateration tech-

nique which estimates the distance between the transmitter and a number of fixed base stations. As shown in Figure 2.3, each distance estimation will geometrically determine a circle, centred at the base station, which indicates the estimated radial distance of the transmitter from the base station. For the RSS method, the estimation of distance is based on an empirically derived propagation loss equation which uses a mathematical model relating propagation loss to distance. Provided there are accurate distance measurements from the transmitter to a minimum of three base stations at known location coordinates, then the transmitters location coordinate can be easily determined. If the distance estimates based on the RSS measurements are larger than the true distances between the transmitter and the receivers a region of uncertainty of the possible transmitter locations is produced. Otherwise, if the estimated distance is smaller than the real distance, three distance measurements may not be able to provide a position fix nor a region of uncertainty. As a result, more than three base stations are normally needed to improve the location accuracy. The weakness in this method is the reliance on a model which may not accurately reflect the channel propagation conditions. Reflection and obstruction by objects, walls or people could cause additional errors in the model.

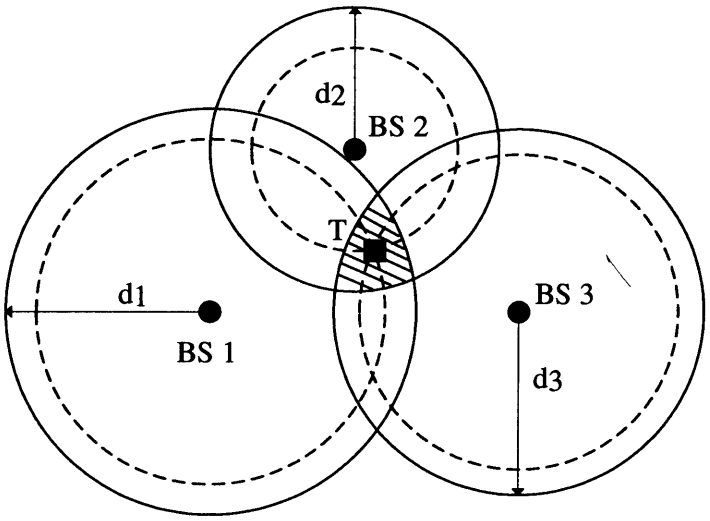


Figure 2.3: Lateration based location estimation.

In TOA the distance from the transmitter to the base station is proportional to the

line-of-sight signal propagation time. The estimated distance from the transmitter to the ith base station is $d_i = ct_i$ where t_i is the signal propagation time and c is the speed of light. The spheres of radius, d_i are seen in Figure 2.3. TOA suffers from two problems. First, all transmitters and receivers in the location system must be precisely synchronised in order accurately identify the LOS component. Secondly, the transmitted signal must include the exact time of transmission so that the base station can discern the distance the signal has traveled. Both requirements are not desirable.

TDOA overcomes the deficiencies in TOA by examining the relative difference in arrival time of the signal between a number of base stations, rather than absolute arrival time at each base station. The TDOA measurement between the two base stations is calculated as a hyperboloid for a constant range difference. The equation of the hyperboloid as Equation 2.1.

$$R_{i,j} = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} - \sqrt{(x_j - x)^2 + (y_j - y)^2 + (z_j - z)^2} \tag{2.1}$$

where: (x_i, y_i, z_i) and (x_j, y_j, z_j) represent the fixed receiver i and j, (x, y, z) is the transmitter co-ordinate position.

If we use three fixed base stations, two hyperbolas ($R_3 - R_1$ and $R_2 - R_1$) are formed from the TDOA measurements. The intersection point locates the estimated position of the transmitter and is shown in Figure 2.4.

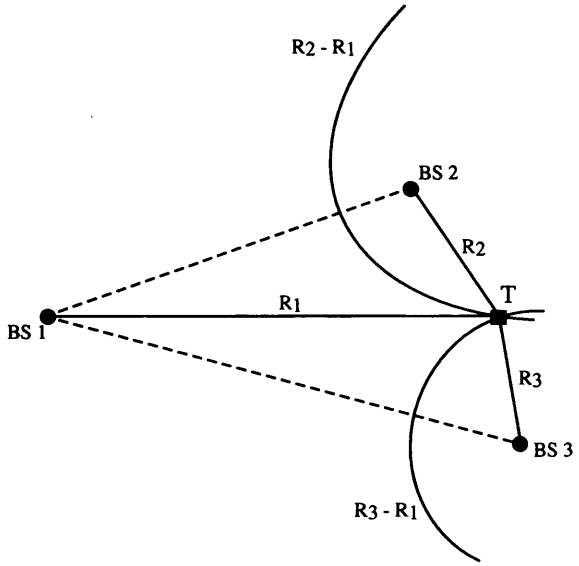


Figure 2.4: TDOA location estimation

The time estimates for TDOA are computed using the cross correlation between the signals received at a pair of base stations. The benefit of using TDOA is that the fixed base stations only require synchronised clocks but this requires one additional base station compared to the TOA technique. However the time delay estimates based on correlation techniques are heavily influenced by the presence of multipath,especially if the reflected rays arrive within a chip period of the direct line of sight (DLOS) ray. This is also true of the TOA technique. To resolve DLOS signal from the NLOS scattering components for a resolution of 1 metre in distance, requires a time resolution of 3 ns or equivalently a 300 MHz pulse bandwidth.

The location metrics have been applied in a number of experimental systems for indoor location. Popular media have included infrared, RF 802.11, and ultrasonic. 802.11 being the most popular choice for indoor location systems since it uses accessible devices that can make use of existing infrastructures. A full and complete taxonomy of all indoor geolocation systems can be found in the following reference [8]. Of particular interest are the location systems that use the RF location metric described in Section 2.2.1. Two examples of the RSS method are LANDMARC: Indoor location sensing using active RFID [9] and RADAR: An in-building RF-based

user location system [10]. LANDMARC is a location sensing prototype system that uses active RFID for locating objects inside buildings. Unlike other systems it employs extra fixed reference tags to help in the RSS location calibration. The experimental results have indicated that when using 4 RF readers arranged in a 4 x 9 m measurement area one reference tag per square meter is required in order to accurately locate objects with an average error of 1 m and a maximum error of 2 m. RADAR uses a radio frequency based system for locating and tracking users inside building with empirical measurements along with signal propagation modelling to determine user location. During the offline phase, signal information was first collected via a 802.11 network of three base stations from a user being tracked around on one floor of a building of size 22.5 m by 43.5 m. The floor was partitioned into many rooms and the user path was along common walkways that ran on the level. The signal strength information was collected at 70 distinct physical locations evenly distributed along the walkway, each location with four directions (North, South East and West). For each combination of location and orientation at least 20 signal strength samples were taken. The mean of the processed signal strength samples was then used to construct and validate models of signal propagation to be used during the online location phase. The experimental results showed, after a suitable choice of an algorithm which matched distance and signal strength, that for a probability of 50 % (median of the CDF of the location error) RADAR is able to estimate a user's location to within 3 metres of the true location. The advantage of the system is the hardware is readily available. However, there are some major disadvantages in the application of the system. The offline measurement set carried out in the paper does not consider the effect of people and represents a single sample set at one period in the day. In addition the signal strength mapping of an entire area has to be repeated whenever the propagation conditions change.

An examination of literature on the AOA technique has shown it has been mainly used in outdoor applications, particularly aircraft location. Reference [7] implements AOA using a 820.11 base station with a revolving directional antenna. The paper proposes the mitigation of propagation effects on RSS range measurements by use

of AOA because RSS methods that rely on range measurements as a function of signal strength, are acutely vulnerable to multipath effects. The experiment involved using a directional antenna situated on a turntable that rotates with a period of 1.8 seconds. The number of signal strength samples taken during the period is 500. The measurement setup used five rotating base stations placed in an area of size 56 m by 25 m. The floor level is partitioned into rooms and the 32 measurement points are evenly spread in the communal walkways within the floor level. For each measurement point, four transmitter orientations (North, South, East and West) were considered. It was found that for each base station for 90% of the time the direction closest to the mobile direction was either the first, or the second highest signal strength peak. After appropriate averaging of the sample bearings, the location error between the base stations at the 50 % percentile in the cdf was between 3 – 4 m. An advantage of the method is that the indoor positioning architecture does not require a signal strength map. The system has several disadvantages. The architecture is mechanical and bulky. This resulted in problems with the angular calibration of the turntable period. A non mechanical implementation should yield comparable performance but this remains unproven. Applying this technique to a system requiring continuous identification and location would mean some tag IDs would not be identified and there would be a tradeoff between direction finding and identification in the system.

There are many types of location metric that could be chosen for the Optag system, each having their own strengths as well as limitations as shown in Table 2.2. Range measurement has been considered, but it is difficult to achieve one metre accuracy without wide bandwidth and complex DSSS techniques. A direction-finding and triangulation solution is preferred because it offers better accuracy combined with *simplicity.* As a result an Angle of arrival (AOA) technique was chosen by the OPTag consortium because this represents a compromise between the complexity of the tag and accuracy of the location system. The location system is implemented using a network of readers, each containing a direction finding array covering the 360° field of view. The antenna elements are squinted at regularly intervals in az-

imuth from each other. The regular arrangement allows the power ratio between adjacent antenna elements to uniquely define the angle of arrival in azimuth. The AOA from at each reader is calculated by the antenna array using differential RSSI from the adjacent receiver channels with the strongest signal strength. The AOAs from a number of readers are then combined at a central location using a suitable algorithm to obtain the estimated location.

Table 2.2: A comparison of the TOA, TDOA, AOA and RSS location techniques

	Accuracy & Resolution	Tag complexity	Min. no. of readers	System cost	Weakness
TOA (two way)	High	High (duplex)	2D:2	High	Synchronisation Large bandwidth
TDOA (one way)	High	Medium	2D:3	High	Requires twice the bandwidth of TOA
AOA	Medium	Low	2D:2/3	Low	Differential RSSI susceptible to multipath
RSS	Low	Low	2D:3	Lowest	path loss model extremely vunerable to multipath
TOA + AOA	High	High	2D:1	V.High	

2.2.2 Channel propagation

An indication of the expected performance of the Optag passive direction finding array prior to the field trials can be found by estimating the received signal strength on each channel of the reader using ray tracing. Ray tracing is a popular method of

predicting the propagation characteristics in mobile systems, particularly in micro-celluar environments. This has the advantage over statistical models, normally used for prediction, in that the signal for each transmitter can be calculated. The ray tracing literature search is focussed primarily on the comparative results between the measured and predicted channel impulse response using indoor propagation modelling. The implementation of a three dimensional ray tracing model for a single floor scenario is described in [11]. The model provides the rapid generation of a complex channel impulse response and is constructed using two techniques, geometrical optics (GO) for surface reflections and the geometrical theory of diffraction (GTD) for diffracted rays off corners or wedges. It assumes that the electromagnetic waves behave as rays so that surface reflections can be found using the theory of ray images. The model is deterministic rather than a ray launching method where the rays are sent out at various angles and their paths are traced until a certain power threshold is reached [12]. The model considers all walls and obstacles as potential reflectors and evaluates the location of their base-station images. The imaging technique works by generating an image table for each base station location, considering all the various wall reflection, transmission, and diffraction permutations within the room. The image information is then used to calculate the channel characteristics at each transmitter location. An additional feature of the model is that a specific illumination zone is calculated for each image. The illumination zone is defined as the area for which the image can give a valid path. This feature significantly reduces the number of images and the time for path tracing and makes it feasible for the model to study large complicated environments with greater speed. A similar method is used in a model presented in [13] in order to produce an efficient algorithm. The model also required the construction of a detailed three-dimensional representation of an example building in the form of a data file which contained the dielectric constant, relative permittivity and thickness for each material. The accuracy of the presented model for a narrowband receiver at 1.8 GHz was then evaluated using comparisons of predictions with measurements in LOS, NLOS, and deep shadow areas, both for co-polarised and cross-polarised antennas (half-wave length dipole). Of particular interest was the accuracy of the predicted average received power. For

the two measurement routes within the building, accurate power predictions can be achieved with rms errors less than 7 dB, even when large sections of the test route are in deep shadow areas. Ray tracing has also been used by [14] in the application of a deterministic ray launching algorithm for the prediction of radio channel characteristics in small-cell environments. It gave a similar theoretical treatment but with greater emphasis on the use of channel sounding and the prediction of rms delay spread of the received signal. These complex models that calculate received signal strength over an infinite number of reflection paths are valuable for gathering information about the propagation characteristics in an indoor environment. However a model limited to a finite number of reflection paths may provide an equally valid insight into the location performance of the DF array, without the need for a complex model that provides a slightly more precise answer but with an order of magnitude increase in complexity and run-time.

3. ACTIVE RF-ID SYSTEM CONSIDERATIONS

The design of the RF-ID system to be prototyped required careful consideration of the main requirements of the system. One of the these requirements was a detection range in excess of 10 m. This range ruled out the use of common place passive near-field RF-ID tags because they operate at very close range, several cm, by inductive coupling in which the tag power is readily derived from the reader unit and communication accomplished by passive modulated back-scattering. A greater operating range to remote reader units, at ranges of typically tens of metres or beyond, was achieved by using an active RF-ID tag containing an internal battery and RF source. The lifetime of the battery within the portable transmitter was then extended by adopting a simplex communication protocol between the tag and reader. The data transmitted by the tag was modulated on a carrier frequency of 5.8 GHz because this allowed the tag antenna to be small as well allowing the transmission frequency to be located in a ISM license free band. The use of this band restricted the transmitter output power to a maximum EIRP (effective radiated isotropic radiated power) of 25 mW. The limit on the transmitter power and the adoption of a simplex communication protocol has implications for the spatial separation of the RF-ID readers and the capacity of each RF-ID cell.

3.1 RF-ID cell size and capacity

The restriction on the transmitter output power, 25 mW EIRP, for the 5.8 GHz ISM band limits the maximum spatial separation of the readers because the operating range is determined by the sensitivity of each receiver. The receiver can correctly identify the tag burst only when the received signal to noise power ratio at the input to the demodulator in the receiver is sufficiently high. The received signal to noise power ratio should be at least 10 dB leading to a bit error probability of no more than 10^{-3} for non-coherent FSK modulation as shown in the waterfall BER plots in Figure 3.1 [15].

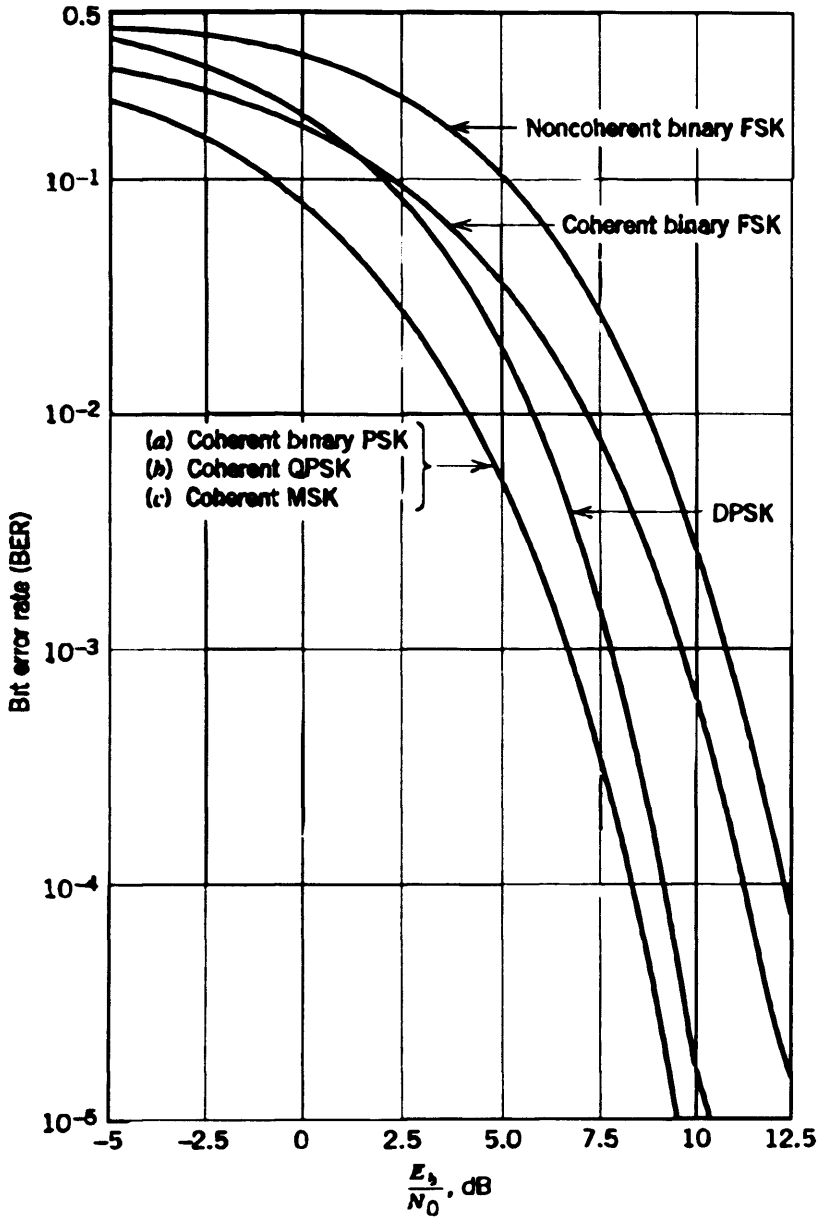


Figure 3.1: Comparison of BER versus Eb/No for different FSK and PSK schemes

For instance at a range of 20 m and an overall receiver noise figure of 4 dB, the received signal to noise ratio whenever the data rate is equal to the filter bandwidth

can be described as follows

$$\frac{S}{N} = \frac{P_t G_t G_r L}{kTBF} = 38.1 \text{ dB} \tag{3.1}$$

where:

- Signal bandwidth, B, of 1 MHz.
- Free space path loss, $L = \frac{\lambda^2}{(4\pi d)^2}$.
- Power transmitted, P_t, of 10 mW.
- Carrier wavelength, λ, of 5.17 cm.
- Line of sight, LOS, distance from tag to reader, d, of 20 m.
- Receiver noise figure, F, of 4 dB.
- Receiver and tag antenna gain, G_r and G_t, of 0 dBi.
- Temperature, T, of 298 K.
- Boltzmann's constant, k, of 1.38×10^{-23} JK^{-1}.

This represents a link margin of 28.1 dB for a desired signal to noise of 10 dB at the input to the demodulator in the receiver. Significantly the link margin excludes the potential affect of fading caused by multipath in a real channel. The type of fading encountered by the receiver is dependent on the relationship between the communication bandwidth and the coherence bandwidth. Table 3.1 represents a listing of some related channel parameters in the transmission path from the tag to the receiver with example values for the additional length of the reflected path and the symbol time within the tag burst.

Since we find that the communication bandwidth is much less than the coherence bandwidth, $W \ll B_c$, the channel between the tag and reader can be considered to be a flat fading channel. All frequencies in the signal are affected equally. Provided there is a significant link margin such as was demonstrated earlier we can assert that

Channel parameter	Symbol	Calculated values
Delay spread	T_d	16.67 ns(for a 5 metre path delay)
Coherence bandwidth	$B_c = \frac{1}{2T_d}$	30 MHz
Symbol time	T_b	1 μs
Communication bandwidth	$W = \frac{1}{T_b}$	1 MHz

Table 3.1: Calculation of the delay spread and coherent bandwidth of the RF-ID system

the receiver provides reliable demodulation and detection of the tag transmission at ranges of 20 m or less even when the transmitted signal is severely affected by multipath and other degradations caused by the channel.

3.1.1 Clashing analysis

The use of a simplex communication protocol between the tag and reader raises concerns about the clashing of data bursts when there are large numbers of tags present in an RF-ID cell. In the prototype RF-ID system, the tag uses pulse repetition interval diversity to avoid repetitive tag clashes with other tags in the same area. The system is asynchronous so that the timing of these data bursts is unrelated and varies randomly from tag to tag. In general if the data burst durations are short in relation to their repetition interval then many tags can be accommodated in the same location with a low probability of error due to data collisions. Since the tags will be transmitting in each interval, it is necessary to know the maximum pulse repetition interval for optimum throughput of tag user IDs.

System definition and analysis: The basic arrangement is shown in Figure 3.2 and consists of a number, n, of far-field RF-ID tags in a single reader cell, each radiating a short data packet of duration, τ, and with pulse repetition interval, T_r. The system is asynchronous so that the timing of these data bursts is unrelated and varies randomly from tag to tag.

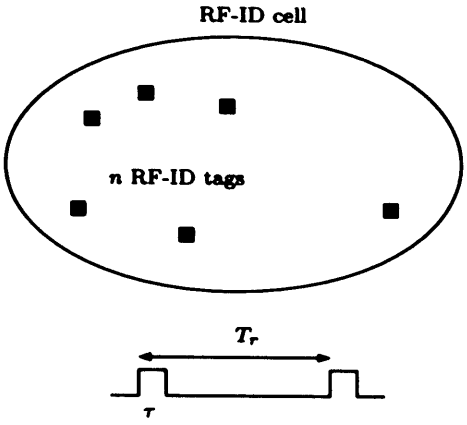


Figure 3.2: A basic arrangement of a far-field RF-ID tag cell

Considering one particular tag sending a single data packet; a collision of some degree will occur if any other tag sends a data packet within $\pm\tau$ s of this transmission. The probability that this data packet will avoid the effects of such collisions from any of the other (n - 1) tags, thus resulting in the survival of the data packet, is therefore given by

$$P(\text{no collision}) = \left(1 - \frac{2\tau}{T_r}\right)^{n-1} \tag{3.2}$$

The mean update interval, over a population of tags, is clearly increased by this finite probability of interference from neighbouring tags, and taking the worst case situation in which all collisions result in loss of the data packet, the mean update interval may be expressed as

$$T_{mean} = \frac{T_r}{\left(1 - \frac{2\tau}{T_r}\right)^{n-1}} \tag{3.3}$$

These results are illustrated in Figures 3.3 and 3.4, for the example of a data packet duration of 150 μs s (150 bits at 1 Mbit/s data rate) and a tag population of between 500 and 2000. It is clear, from Figure 3.3, that the probability that data collisions are avoided increases progressively with the repetition interval and with smaller tag populations; however, from Figure 3.4, an optimum repetition interval

may be chosen to give minimum mean update interval and hence maximum data throughput.

Optimum data parameters: The usual situation is for the data packets to be of fixed duration and the tag population to be a known maximum value. For this case, the repetition interval may be optimised for minimum mean update interval, as follows

$$\frac{\delta T_{mean}}{\delta T} = \frac{\left(1-\frac{2\tau}{T}\right)^{n-1} - \frac{2\tau}{T}(n-1)\left(1-\frac{2\tau}{T}\right)^{n-2}}{\left(1-\frac{2\tau}{T}\right)^{2n-2}} = 0 \text{ at } T = T_{opt} \tag{3.4}$$

which results in

$$T_{opt} = 2n\tau. \tag{3.5}$$

The minimum mean update interval can be then found by the substitution of T_{opt} as follows

$$T_{mean}\,(\min) = \frac{2n\tau}{\left(1-\frac{1}{n}\right)^{n-1}} \rightarrow 2en\tau \equiv eT_{opt} \text{ for large } n \tag{3.6}$$

Taking again the illustrative values used Figures 3.3 and 3.4, for a tag population of 1000, the optimum choice of repetition interval is thus 0.3 s and the corresponding minimum update interval is 0.82 s. For an optimised repetition interval such as this, by substituting T_{opt} into Equation 3.2, the probability that a given data packet suffers no collision is

$$P\,(\text{no collision optimised}) = \left(1-\frac{1}{n}\right)^{n-1} \rightarrow \frac{1}{e} \tag{3.7}$$

and so maximum data throughput requires that some 63% of data packets are lost due to collisions. Another interesting observation to note from the analysis is if one operates the asynchronous simplex data protocol at the optimal repetition rate, N_{opt}, described as

$$N_{opt} = \frac{T}{T_{opt}} = \frac{T}{2n\tau} \tag{3.8}$$

one is allotting each tag exactly half of the time allotment T/N available in a synchronous system.

Conclusions: An analytic technique has been presented that allows prediction of the data packet collision probability and its effect on data throughput of far-field RF-ID tags using an asynchronous simplex data protocol. This method does not make most efficient use of the radio channel (since only 1/e of all tag bursts are received in the optimal case and the tags transmit, each, half as often as in the synchronous case), but leads to an extremely simple hardware implementation for the individual tags in the RF-ID system. No synchronization information at all has to be received from the reader; therefore, a one-way communication from the RF-ID tags to the reader is sufficient.

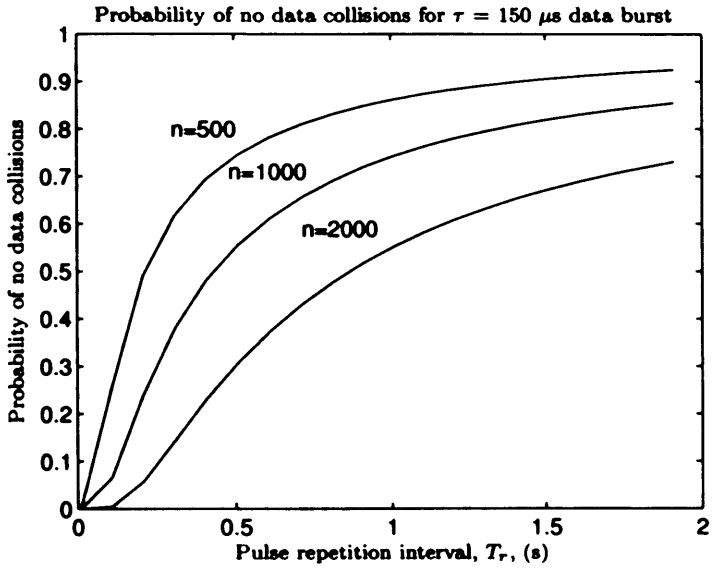


Figure 3.3: Data throughput - Data packet survival probability versus repetition interval and tag population

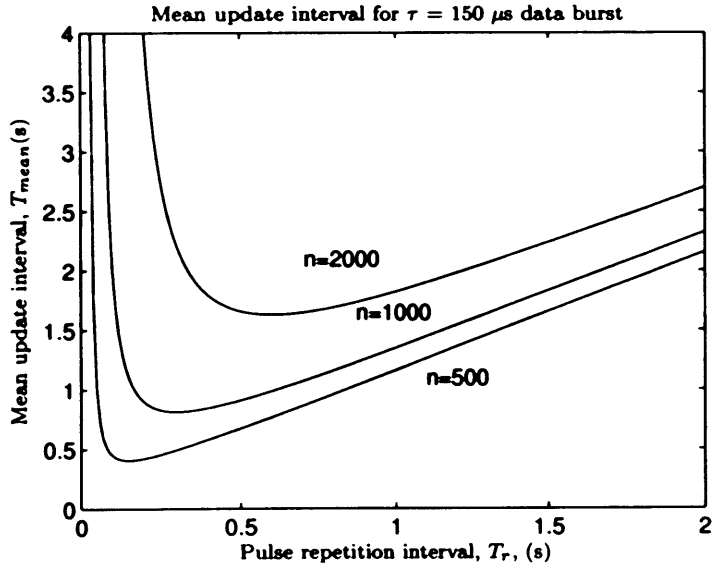


Figure 3.4: Data throughput - Mean tag data update interval versus repetition interval and tag population

3.2 Investigation of an RF-ID receiver incorporating frequency diversity

Whilst pulse repetition interval diversity reduces the potential for clashing between a number of tags transmitting on the same carrier frequency, it does not prevent in band frequency interference from other systems operating in the ISM band from disturbing the narrow band signal transmitted from the tag to receiver, particularly if the interference is continuous in nature. This form of interference could be avoided by allowing the tag to transmit from burst to burst at carrier frequencies anywhere within a band centred on 5.8 GHz rather than at a fixed frequency. The size of the band determined by the detection agilty of the receiver and the chosen frequency stability of the frequency source within the transmitter. Besides making the RF-ID system less susceptible to frequency interference, oscillators of low frequency stability are less expensive than stable fixed frequency PLL sources and therefore provide a lower cost if the RF-ID tag was embedded for example in a boarding card in an airport. By adopting this approach the complexity of the system is now referred to the receiver.

In the proposed system to be investigated, an active RF-ID tag transmits a BPSK (Binary phase shift keyed) modulated carrier to a receiver. The tag can transmit at carrier frequencies anywhere within a wide band of 30 MHz centred on 5.8 GHz but in principle an arbitrarily-wide band can be accommodated by the receiver. The synchronisation technique in the receiver estimates the carrier frequency using the spectral properties of the received pre-amble in the data burst. The properties of the frequency estimator are now examined prior to introducing the architecture of the receiver system.

3.2.1 Estimation theory

Frequency measurement and acquisition can be examined using the following two areas: detection theory [16] and sinusoidal parameter estimation [17]. The analysis is evaluated under the assumption that the modulated carrier signal at the receiver can be considered to be stationary and the modulation components have sufficient

separation as to be detected. An unbiased estimator provides a best estimation of the parameters given the available data observations. Maximum likelihood estimators for single tone and multiple tone parameter estimation from discrete observations are detailed in [17, 18]. Equation (3.9) shows the maximum likelihood estimator, X_{ML}.

$$X_{ML} = \max_f \left| \frac{1}{N} \sum_{n=0}^{N-1} x_n \exp\left(-j2\pi f n T_s\right) \right|^2 \tag{3.9}$$

where: x_n for $n = 0, 1.., N-1$ are the available input samples and T_s is the sampling time. The index with the maximum amplitude is the chosen frequency.

The Cramer-Rao bound describes the lower bound on the variance of any unbiased estimator. Equation (3.10) shows the Cramer-Rao bound for the maximum likelihood estimator provided there is a single real sinusoidal signal embedded in white, Gaussian noise, with unknown frequency, unknown amplitude and unknown phase:

$$\sigma_f = \frac{\sqrt{12}}{2\pi T_s \sqrt{N\left(N^2-1\right) SNR}} \tag{3.10}$$

where: N is the number of samples and SNR is the signal to noise power ratio.

Indeed the Cramer-Rao bound accurately describes the frequency error in the ML estimator above a certain signal to noise threshold. Since the maximum likelihood estimator is very similar to a Fourier transform, a possible method for coarse frequency estimation of a single frequency is the FFT. This provides an estimate evenly spread over a set of frequency bins based on the number of samples in the transform. The frequency resolution, f_{res}, of the FFT is inversely proportional to the data time length, T_w, as shown in (3.11). It can also be expressed in terms of the sampling time, T_s, and the number of samples in the transform, N, which are taken from the time domain signal.

$$f_{res} = \frac{1}{T_w} = \frac{1}{NT_s} \tag{3.11}$$

In fact, the FFT estimator can also provide an approximation to the maximum likelihood estimator for multiple frequencies separated by more than $\frac{4}{NT_s}$ or four bins as derived by [18]. The resolution of the FFT estimator can be improved by adding zeros to the end of the signal. This achieves a longer FFT length without modifying the spectral content of signal. Because the zero-padded signal is long (though no new energy has been added), the resulting FFT provides a better frequency resolution. The resolution also determines the maximum error expected from the estimator which occurs whenever when a tone is located halfway between bins. If the tone frequencies are uniformly distributed in $\pm\frac{f_{res}}{2}$ then the standard deviation of the frequency estimate from a non-zero padded FFT is

$$\sigma_{coarse} = \frac{1}{2\sqrt{3}NT_s} \tag{3.12}$$

[19] highlights an improvement in the FFT accuracy by using fine frequency estimation and applying windowing functions to achieve maximum dynamic range in the resolution bin. However an optimum FFT-based frequency acquisition may not be required for this application because as long as a carrier frequency estimate is estimated within the capture bandwidth of the demodulator and before the end of the pre-amble of the tag data burst, thus there is no need to perform fine frequency estimation to overcome truncation from the FFT bins.

Now that frequency measurement has been considered, previous attempts at frequency acquisition must be evaulated prior to incorporating carrier synchronisation into the receiver design. A considerable amount of literature has been published on carrier synchronization in the presence of small frequency offsets. These studies are applied to applications where the frequency offset varies by no more than a few 100 kHz. For instance, [20] highlights a method for coarse frequency acquisition for Nyquist filtered MPSK. The estimation procedure looks for a small frequency offset when a BPSK modulated carrier is downconverted to a low pass filtered baseband signal. The offset in the carrier is estimated from the spectral components in the bandlimited BPSK signal. In another example, [21] examines the use of a DFT-based frequency acquisition algorithm for carrier offsets in mobile satellite re-

ceivers. An open loop DFT estimation method along with feed forward frequency estimation algorithm is demonstrated with a capture range ΔF of only ± 5 kHz. These systems operate on very small variations in the carrier frequency. A more sophisticated receiver design must be considered for much larger variations in the carrier frequency.

3.2.2 Receiver system

The architecture shown in Fig. 3.5 uses the spatial and frequency domains to detect and estimate the carrier frequency of a short pre-amble burst from each tag transmission. The estimated frequency from the carrier frequency estimator (CFE) module is then used to set the frequency of a direct digital synthesiser (DDS). The signal is then down converted to a well-defined intermediate frequency (IF) for demodulation by a Costas loop. The design also contains an AGC the design of which must be carefully considered since it must not affect the sensitivity of the carrier frequency estimator (CFE) which performs a two dimensional search in frequency and amplitude during the duration of the tag pre-amble. The carrier frequency is estimated using the arithmetic mean of the primary modulation indices identified during the duration of the tag pre-amble. The tag pre-amble consists of alternate zeros and ones modulated onto the carrier using Binary phase shift keyed (BPSK) modulation. As a consequence, the pre-amble contains harmonics that have frequencies at odd multiples of, $R/2$ Hz where R is the data rate. The spectral pair with the largest amplitude are the primary modulation components which are separated by R Hz. These are then used to estimate the carrier frequency that sets the frequency of the DDS.

By way of an example consider a tag data burst that is received with a random carrier frequency, f_c within the detection bandwidth, Δf_b, centred on a convenient first intermediate frequency stage. The tag signal bandwidth, Δf_s, is chosen as 1 MHz which means the random carrier frequency must be resolved to a frequency resolution, f_{res}, of 125 kHz (here a resolution of 1/8 of the data rate, R, has been chosen). To accommodate this frequency resolution the pre-amble of the received signal must be measured for 8 μs. The duration of the tag pre-amble, T_w, is usually

set greater than 8 μs to account for any latency encountered in the synchronisation procedure. A simple peak search algorithm is then able to detect the primary modulation components and distinguish them from even high-level noise or other interference, thus estimating its frequency to a precision of $\pm$ 62.5 kHz. The DDS is then set to the estimated frequency, $f_c \pm f_{res}$ but with an offset equal to the desired final IF frequency, $f_c \pm f_{res} - f_{IF}$ and then down-converts the signal to its final IF frequency, f_{IF}. Since the signal bandwidth, Δf_s, is around 1 MHz it is now an easy matter to demodulate by standard techniques, such as a Costas loop, which will acquire lock within a further couple of bit periods, well before the end of the pre-amble. For a $\frac{c}{n_o}$ of 2×10^7 (quite low), and an acceptable phase jitter of 7 ° rms, the loop natural frequency of the Costas loop should be 100 kHz. This should give fast tuning (around 2 μs) of the Costas loop, provided the initial frequency offset is no more than around 200 kHz - say 2 FFT resolution cells [22].

For the 30 MHz detection bandwidth, Δf_b, considered in this design and an intermediate first stage frequency of 100 MHz, a 80 MHz bandpass sampler is sufficient in an FFT implementation of the carrier frequency estimator. The corresponding FFT size, N, can be calculated in the following manner. By rearranging (3.11) the FFT size is equal to $\frac{T_w}{T_s}$. Since the sample time, T_s, from the 80 MHz ADC is equal to 12.5 ns and T_w was calculated as 8 μs, the FFT size, N is equal to 640 points. A high end dedicated DSP chip or an FPGA with a DSP core could perform an FFT of this size with a short latency of a few microseconds. The disadvantage of both approaches is the high purchase cost and additionally for the FPGA chip the development of a multilayer PCB board for the BGA chip interface. A viable alternative to an all digital CFE is an analog method that estimates the Fourier transform using surface acoustic wave (SAW) linear dispersive delay lines. This approach also removes the need for a 80 MHz ADC used in the digital implementations and the sliding Fourier transform estimate can be performed at a relatively high intermediate frequencies.

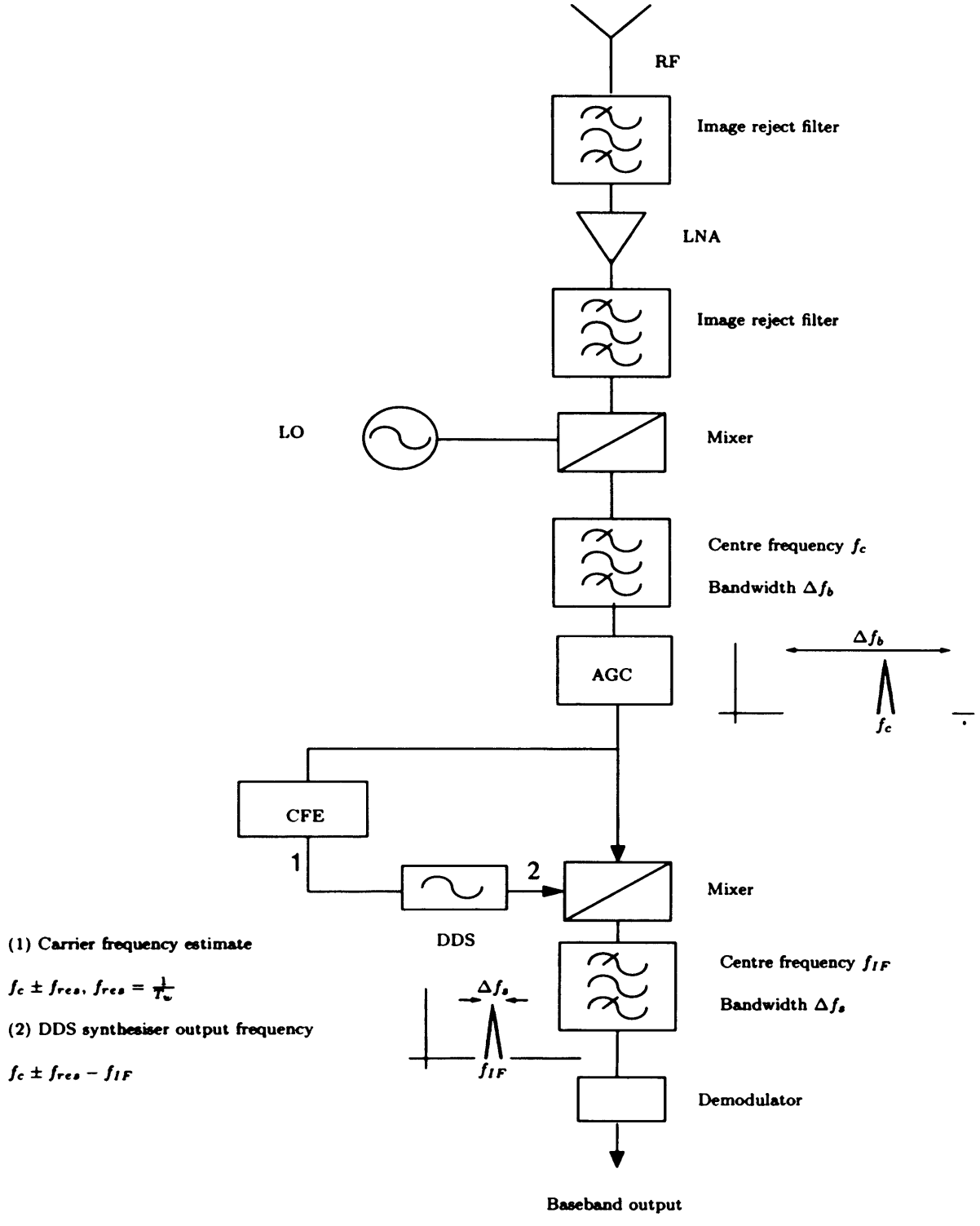


Figure 3.5: Tag reader receiver architecture

3.2.3 Implementation of the carrier frequency estimator (CFE)

Chirp filters are commonly used in pulse compression radar systems. A second application of these devices is the compressive receiver which is a system for frequency measurement [23]. This system takes advantage of the linear dispersion delay with frequency property of the chirp filter in the measurement of the Fourier transform of the received signal. The chirp filters, which are also known as linear dispersion delay lines, have an impulse response with an instantaneous frequency that varies linearly as a function of time. The impulse response of the chirp filter can be shown for a constant chirp rate $\mu(t)$ as follows

$$g(t) = a(t)\cos(w_c t + \pi\mu t^2 + \phi_o) \tag{3.13}$$

where the chirp signal has length T_c, centre frequency w_c at time $t = 0$ so that $a(t) = 0$ for $|t| > \frac{T_c}{2}$, the envelope is flat so that $a(t)$ is constant for $|t| \leq \frac{T_c}{2}$ and ϕ_o is a constant.

This response is utilised in the SAW chirp-Z-transform to form the power spectrum of the received signal,

$$F(2\pi\mu\tau^2) = e^{-j\pi\mu\tau^2} \int_{-\infty}^{+\infty} \left(x(t)e^{j\pi\mu t^2}\right) e^{j2\pi\mu(\tau-t)^2} \tag{3.14}$$

The process can be described as the multiplication of the input signal by a chirp waveform, followed by convolution with a chirp filter; the post multiplication with a third chirp waveform is used to correct the phase of the output waveform and is not needed if the power spectrum is desired. The power spectrum can therefore be found using a pair of chirp filters with opposing chirp directions. This method of frequency analysis is used quite widely in radio astronomy and microwave remote sensing [24, 25].

A schematic of the compressive receiver is shown in a M(long)-C(short) scheme in Figure 3.6. The impulse response of the expander is longer than the compressor, to

ensure that the bandwidth of the signal can be completely compressed and provide an accurate spectrum estimate. As before the input signal x(t) is mixed with a linear downchirp waveform (expander) followed directly by a convolution in a linear upchirp filter (compressor). It can be shown that the output signal v(t) is a sliding Fourier transform [23].

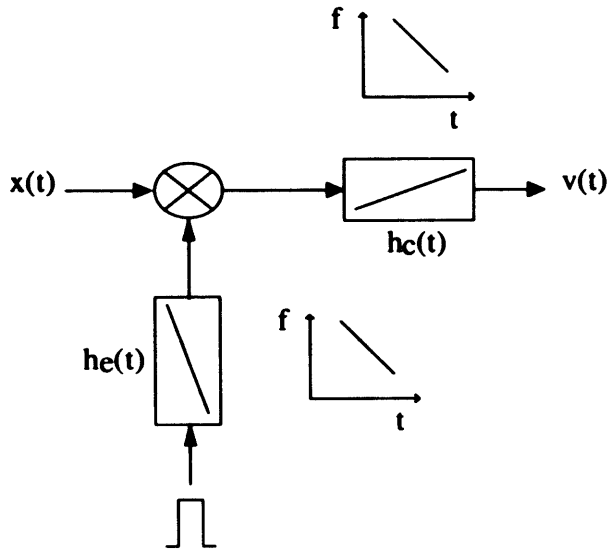


Figure 3.6: Schematic of the compressive receiver

Architecture of the CFE: The duty cycle is an important parameter when considering the compressive receiver as a frequency estimator because a carrier frequency estimate is required after each pre-amble duration, T_w, so that the DDS can be set to the correct frequency to allow the Costas loop to demodulate the data transmission. The repetition time of the CFE module cannot be shorter than the compressor duration to avoid time-frequency ambiguity. One module will therefore only have a 50 % duty cycle and so two modules are required in push-pull mode to obtain a 100 % duty cycle as seen in the SAW CFE architecture shown in Figure 3.7. In this architecture, the input signal is passed to both Fourier transform modules operating in alternate fashion. The modules are switched at the appropriate times using the timing generator and RF switches. In each cycle the input signal is multiplied in a

linear mixer with a linear expanding chirp waveform with a flat spectrum. This flatness is achieved with a saturating amplifier. On post multiplication the convolution is followed by suitable IF amplification. The power spectrum is assessed using a square law detector via additional low pass filtering and amplification. The output of the detector is sampled throughout the T_w period by an ADC and these samples are sent to a microcontroller. If there is sufficient time these samples could then be passed to an acquistion and integration circuit to improve the accuracy of the amplitudes estimates.

The parameter bounds of the time dispersion, T_e, T_c, and the bandwidth, B_e, B_c, of the SAW chirp filters for the design are chosen to be well within those described by [26]. In general TB (time-bandwidth) products up to about 10,000 are possible. In this example, the compressor centre frequency is chosen as 100 MHz and a corresponding bandwidth, B_c, of 30 MHz. Working backwards, the expander must then mix down the RX signal to a band centred at the compressor centre frequency. In addition to ensure the frequency band of the compressor impulse response completely overlaps the signal band the expander bandwidth, B_e, must be chosen to be twice the compressor bandwidth, B_c, because the literature shows an optimum expander/compressor chirp length ratio of 2 to achieve a rectangular transform window [25, 26]. Assuming the input RX signal is chosen for convenience at a common IF of 200 MHz and has a bandwidth of $\pm$ 15 MHz then the expander centre frequency must therefore be chosen as 100 MHz with a bandwidth, B_e, of 60 MHz. The frequency resolution of the system is equal to the reciprocal of the length of the compressor impulse response, T_c. Hence for a desired frequency resolution, f_{res}, of 125 kHz (so as to be well within the Costas Loop capture range), T_c, must be equal to 8 μs. This is equivalent to 240 resolvable points - either by $\frac{T_c B_c}{4}$ or alternately the compressor bandwidth B_c divided by the desired frequency resolution, f_{res}. Consequently if we retain the same chirp rate μ for both chirp filters the expander impulse response, T_e, is equal to 16 μs.

Conclusions: The dynamic range of the CFE will be reduced by not taking account

of side lobe weighting to remove spectral leakage as a result of the finite Fourier transform estimation window, T_c. The performance for non-bin centre frequencies could be improved by as much as 0.82 dB using a Hamming window method. However at high SNR there should only be a small effect on the quality of the power spectrum estimate. An additional source of error is manufacturing tolerances. Quartz material which is used to make the filters have a surface wave velocity which can be controlled to within 50 parts per million or better. At 100 MHz this is only 5 kHz error per tag transmission. This represents a negligible amount of error.

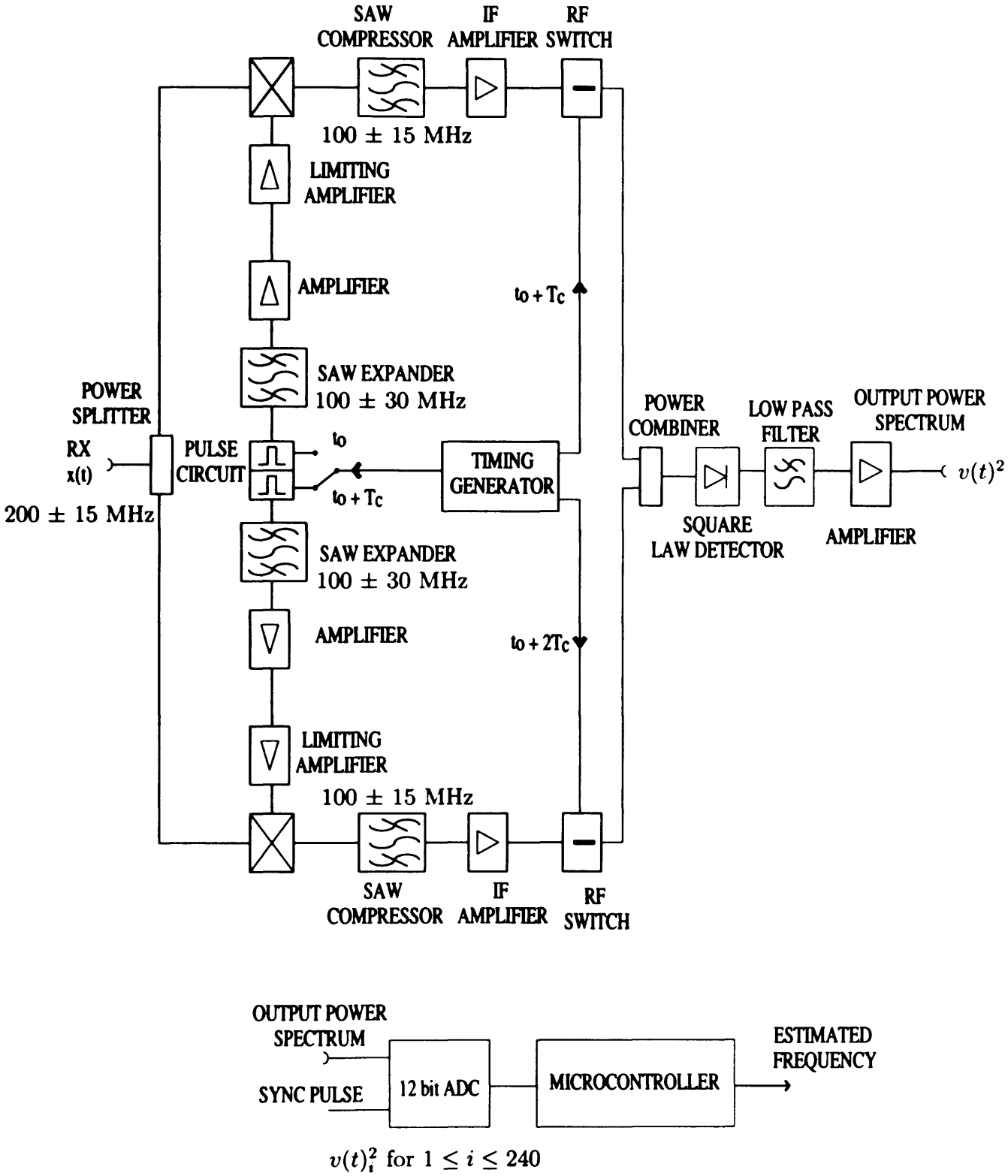


Figure 3.7: Architecture of the SAW carrier frequency estimator

3.2.4 Analysing the performance of the carrier frequency estimator

Before analysing the performance of the two peak search algorithms used in the carrier frequency estimator we must model the received signal using the discrete FFT implementation in which the carrier signal is downconverted from a 100 ±15 MHz band down to a 20 ±15 MHz frequency band as a result of undersampling by the ADC.

BPSK model: We begin the analysis of the estimator by showing that the baseband BPSK signal can be written as

$$S(t) = \sum_{n=-\infty}^{\infty} a_n p(t - T_b) \tag{3.15}$$

where: $p(t)$ is the baseband NRZ pulse limited to T_b seconds and a_n is a preamble periodic sequence of alternating ones and minus ones representing binary ones and zeros.

Applying Fourier analysis to $S(t)$ using an FFT of size, N, we find that frequency content in each FFT bin, $S(k)$, shown in Equation 3.16 is a weighted sum of harmonics that have frequencies at odd multiples of $\frac{1}{2T_b}$ Hz.

$$S(k) = \sum_{n=0}^{N-1} S(n) exp(\frac{-j2\pi kn}{N}) = A \left(\frac{\sin\left(\frac{\pi(f-f_c)}{R}\right)}{\left(\frac{\pi(f-f_c)}{R}\right)} \right) \tag{3.16}$$

The amplitude and separation of the two spectral pairs are illustrated in Figure 3.8 whenever the BPSK binary modulation is superimposed on a carrier signal.

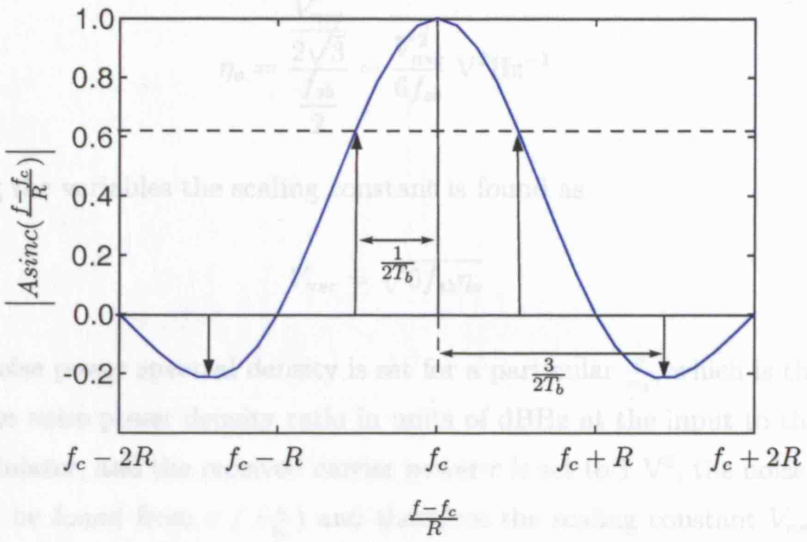


Figure 3.8: Theoretical spectra for a BPSK modulated carrier

Setting the noise in each FFT bin: Since the signal is located somewhere within the 5 - 35 MHz frequency band as a result of undersampling by the ADC, the output of undersampling by the ADC can modelled as

$$s(t) = Am(t)\cos 2\pi f_c t + \eta(t) \quad 0 \leq t \leq T_b \tag{3.17}$$

where: A is an amplitude constant, $m(t) = +1$ or -1, f_c is the unknown carrier frequency, T_b is the symbol duration of 1 μs, $\eta(t)$ is a white Gaussian noise process with a mean of zero in a sampling bandwidth, $\frac{f_{sb}}{2}$, of 40 MHz.

The signal to noise power ratio of the signal received at the Costas Loop can be set by appropriate voltage scaling of a white Gaussian noise process. The rms voltage of the noise process is presented in the analysis in Appendix A.1 and can be given by

$$\frac{V_{nsc}}{2\sqrt{3}}$$

and can be adjusted using the scaling constant V_{nsc}. The noise spectral density in

the sampling bandwidth can also be described as

$$\eta_o = \frac{\frac{V_{nsc}}{2\sqrt{3}}}{\frac{f_{sb}}{2}} = \frac{V_{nsc}^2}{6f_{sb}} \text{ V}^2\text{Hz}^{-1} \tag{3.18}$$

rearranging the variables the scaling constant is found as

$$V_{nsc} = \sqrt{6f_{sb}\eta_o} \tag{3.19}$$

Since the noise power spectral density is set for a particular $\frac{c}{n_o}$, which is the carrier power to the noise power density ratio in units of dBHz at the input to the Costas loop demodulator, and the received carrier power c is set to 1 V^2, the noise spectral density can be found from c / $(\frac{c}{\eta_o})$ and therefore the scaling constant V_{nsc} can be set in the model.

Having described a method of setting the rms noise level in each FFT bin relative to a normalised received carrier power, c, of 1 V^2 we must then determine the absolute levels of the primary and secondary modulation spectral pair within the signal. First using (3.17) the amplitude, A, of the carrier is known to be equal to $\sqrt{2c} = 1.41$. The amplitude of the sidebands can then be calculated using (3.16), the first pair of sidebands have amplitude denoted $0.636A$ and frequency separation $\frac{1}{T_b}$ Hz; the second pair of sidebands have amplitude denoted $0.212A$ and frequency separation $\frac{3}{T_b}$ Hz. Figures 3.10 and 3.9 illustrate an FFT of the BPSK modulated carrier with added noise. The noise is calculated for a desired signal to noise power ratio of 20 dB and 7 dB in 1 MHz signal bandwidth.

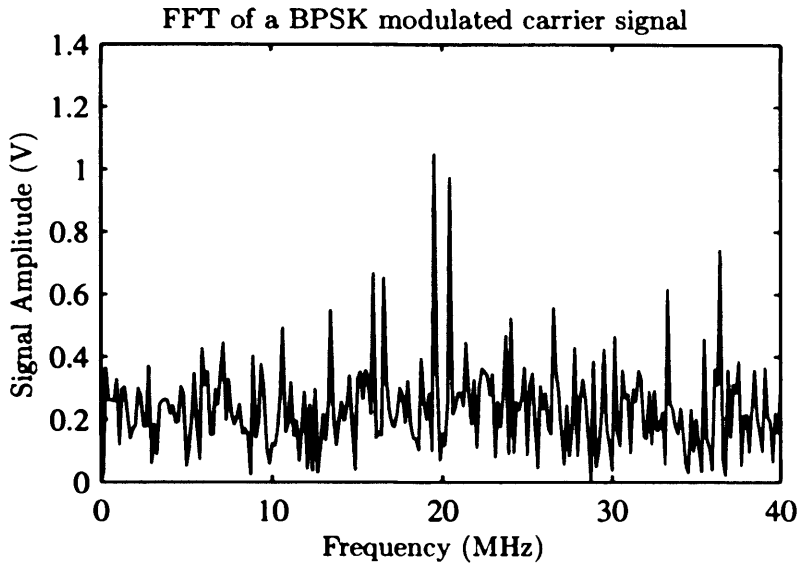


Figure 3.9: FFT of a BPSK modulated carrier signal with $\frac{c}{n_o} = 5 \times 10^6$ Hz (SNR = 7 dB)

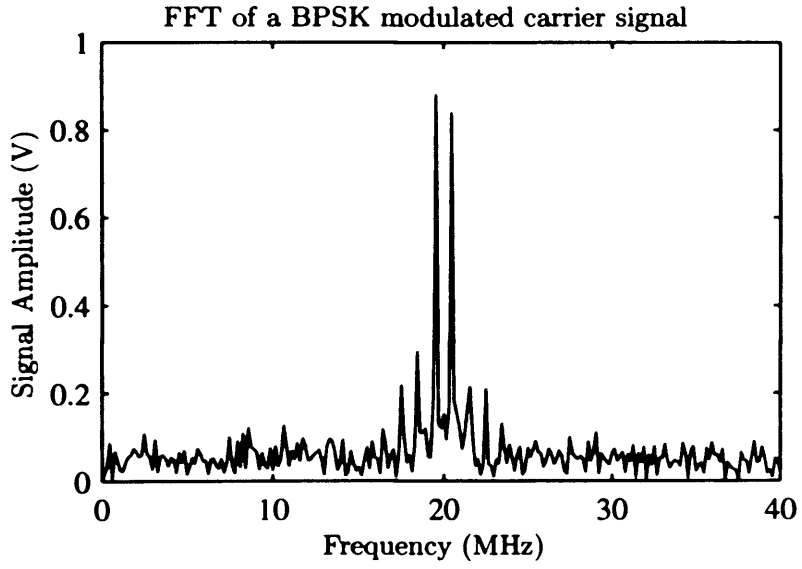


Figure 3.10: FFT of a BPSK modulated carrier signal with $\frac{c}{n_o} = 100 \times 10^6$ Hz (SNR = 20 dB)

Frequency estimation considerations: A number of factors must be evaluated before considering the accuracy of any potential peak search algorithm. The larger the number of cycles of data that is processed by the FFT then the greater is the likelihood of a successful peak search at low signal to noise power ratios. Since we can never have access to an infinite data record, a short time window must then be used to compute the FFT. This can lead to spectral leakage if the FFT is performed over a non-integer number of cycles of the input frequencies. This could lead to false detections by algorithms which use the two highest peaks as representing the primary spectral pair ; for example if the two amplitudes reflect the same modulation position i.e. two samples on the same crest. Another problem occurs whenever the data window is not a power of two, in which case the FFT algorithm will append zeros to the data for the desired FFT window size and the frequency impulses in the spectrum will become a non-desirable sinc shaped. This can be explained by imagining the addition of d zeros is processed by the FFT in similar manner to taking a sinusoid and multiplying it with a rectangular box of length d. The multiplication of the box and a sinusoid in the time domain results in the convolution of a sinc with impulses in the frequency domain.

In light of the above discussion, the data is processed by the carrier frequency estimator in the following manner in order to obtain the best possible fit between the modeled and the simulated results for the peak search algorithms. First, the FFT window is maximized to obtain as high a resolution as possible considering technology, time constraints and the capture range of the demodulator. Second, the time window, T_w, is sampled with an FFT size, N, which is a power of 2. This results in no appended zeros and therefore no sidelobes due to 'sincing'. Third, the number of symbol cycles, N_{cycles}, which is calculated from $\frac{T_w}{T_b}$, is chosen to be an integer value so that there will be no spectral leakage. Provided the rules are followed the simulated and modeled results of the peak search algorithms discussed next should match consistently.

Analysis of peak search algorithm: The performance of the two peak search algorithms will be measured using the probability that the FFT search algorithm fails to estimate the carrier frequency with error no greater than the FFT resolution. The analysis of the algorithms will be based on the assumption that the received carrier frequency and the modulation components are separated by even multiples of the frequency resolution, i.e. on a bin centre. This is a first order approximation to a real situation where the carrier and its modulation components can be non-bin centred. In practice there is a fixed error associated with randomly received signal frequencies because they can be situated anywhere within the resolution bandwidth of $\pm\frac{f_{res}}{2}$. For instance if the carrier and its modulation components are located midway between bins, the magnitude of the FFT estimation at each of the bins drops to 0.64 of its maximum value and this reduces the probability of detection of the primary modulation components at low SNR.

3.2.5 Algorithm I

This algorithm takes the average index value from the two indices $k_{i=1,2}$, corresponding to the two largest peaks in the FFT of the received signal, $S(k)$, with the magnitude described as

$$P_k = |S(k)|, k = 0 \text{ to } N \tag{3.20}$$

and forms a frequency estimate $\hat{f}_c$ of the carrier frequency f_c. It may be examined experimentally by simulation and also by analysis which extends the work of [17].

The output of the FFT for a signal containing modulation harmonics and noise can be modeled in the following manner. The magnitude P_k is a random variable whenever there are signals and noise present. In this analysis we neglect the condition of the signal frequency being half the sampling frequency for zero bias. The noise samples remain independent, normal and have zero mean with variance σ^2. Each non-tone bin can be shown to be a Rayleigh distribution denoted by a distribution $C_n(P_k)$ and each tone bin has a Rician distribution ; for the first spectral pair it is denoted $C_n(P_{r1})$ for tones with indices $k = r1_a, r1_b$, and for the second spectral pair it is denoted $C_n(P_{r2})$ for tones with indices $k = r2_a, r2_b$. The other

modulation components are neglected and are modeled as non-tone bins.

$$C_{m1}(P_{r1}) = \frac{P_{r1}}{\sigma^2} \exp\left(\frac{-(P_{r1}^2 + b^2)}{2\sigma^2}\right) \cdot I_0\left(\frac{bP_{r1}}{\sigma^2}\right), \quad P_{r1} \geq 0 \tag{3.21}$$

$$C_{m2}(P_{r2}) = \frac{P_{r2}}{\sigma^2} \exp\left(\frac{-(P_{r2}^2 + c^2)}{2\sigma^2}\right) \cdot I_0\left(\frac{cP_{r2}}{\sigma^2}\right), \quad P_{r2} \geq 0 \tag{3.22}$$

$$C_n(P_k) = \frac{P_k}{\sigma^2} \exp\left(\frac{-P_k^2}{2\sigma^2}\right), \quad P_k \geq 0, \quad k \neq r1_a, r1_b, r2_a, r2_b \tag{3.23}$$

where: σ is the rms noise power in each FFT bin, b is the peak primary modulation amplitude, and c is the peak secondary modulation amplitude

Having defined the probability distributions we can now develop the analytic probability of failure for Algorithm I. The probability that one of the two primary tone indices $r1_a$ $r1_b$ has the largest amplitude among all other bins is expressed as

$$\mathrm{P_s} = [\Pr\{\text{all } C_n < C_{m1}\} \cap \Pr\{\text{both } C_{m2} < C_{m1}\}] \tag{3.24}$$

This be further expanded as

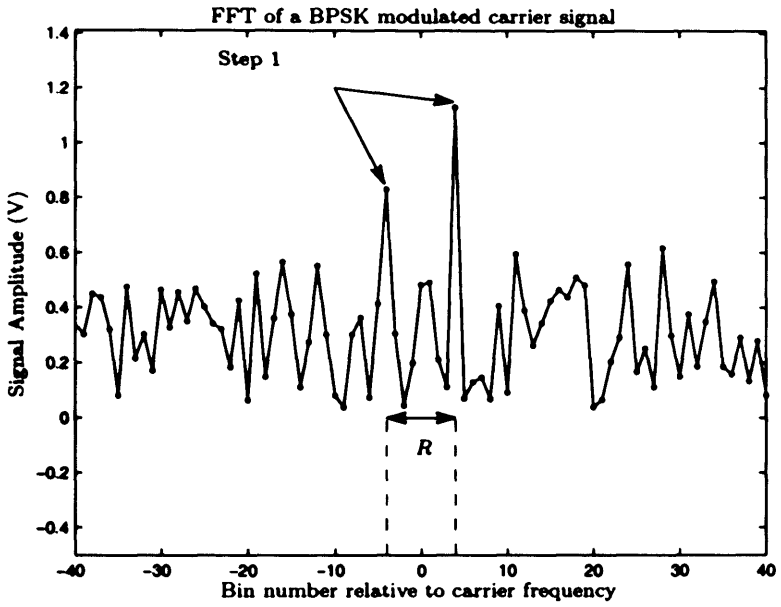
$$\mathrm{P_s} = \int_x \Pr\{\text{all } C_n < C_{m1} | C_{m1} = x\} \cdot \left[\int_0^x \Pr\{\text{both } C_{m2} < C_{m1} | C_{m1} = y\}\, dy\right] \cdot \Pr\{C_{m1} = x\}\, dx \tag{3.25}$$

This can be expanded by the multiplicative probability of the independent pdf distributions within each FFT bin. This is formed using the cdf of the Rayleigh distributions for the non-tone bins – that can be expressed analytically and the cdf of the Rician distributions for the second pair of spectral components – that can

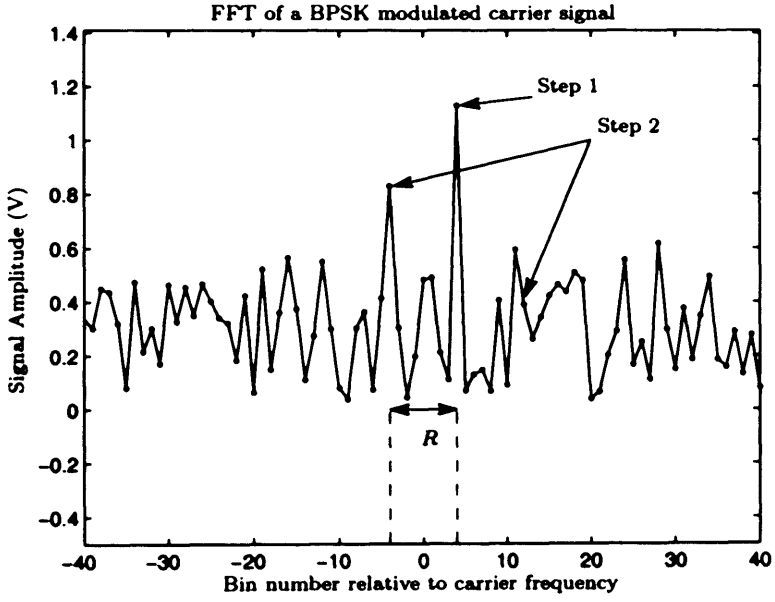
only be expressed numerically. Equation (3.26) shows the final evaluated function with substituted distributions pdf and cdf .

$$
\begin{aligned}
P_s \;=\; & \int_0^{\infty} \left[1 - \exp\left(\frac{-x^2}{2\sigma^2}\right)\right]^{N-4} \\
& \cdot \left[\int_0^{x} \frac{y}{\sigma^2} \exp\left(\frac{-(y^2+c^2)}{2\sigma^2}\right) I_0\left(\frac{cy}{\sigma^2}\right) dy\right]^2 \\
& \cdot \left[\frac{x}{\sigma^2} \exp\left(\frac{-(x^2+b^2)}{2\sigma^2}\right) I_0\left(\frac{bx}{\sigma^2}\right)\right] dx
\end{aligned}
\tag{3.26}
$$

The probability of failure is the probability that two primary tone bins fail to be the highest two values. Since both primary tone bins have probability distributions which are independent, the probability of both being the highest two values is $P_s{}^2$. Therefore the probability of FFT failure is $P_{\text{failure}} = 1 - P_s{}^2$. The simulation and analysis were implemented using Matlab as shown in Appendix A.2. The accuracy of the analysis is demonstrated first for the probability of failure for a FFT of size 512 when compared to a simulation with 3000 trials as shown in Figure 3.12. The accuracy of the model is repeatable for other FFT sizes as shown in Figures 3.13 and 3.14 which show that the modeled and simulated results for algorithm I closely match each other.



(a) Algorithm I - (1) Pick the highest two peaks



(b) Algorithm II - (1) Pick the highest peak and then (2) Pick the highest peak in a 2R Hz band centred on the peak chosen in (1)

Figure 3.11: Illustration of the peak-search algorithms for a BPSK modulated carrier signal with $\frac{c}{n_o} = 3 \times 10^6$

3.2.6 Algorithm II

This algorithm takes the largest spectral peak with index, $k_{i=1}$, and assumes it to be one of the primary tones. It then takes the second largest spectral peak with index, $k_{i=2}$ from a 2R Hz bandwidth centred on the index of the first spectral peak and then assumes this to be the second primary tone. The algorithm is successful provided the first spectral peak happens to coincide with one of the primary tone indices and the other primary tone has the highest amplitude in a bandwidth 2R centred on the index of the first spectral peak. The analysis of the algorithm can be approximated by the following method. The probability of FFT success is the probability that one of the two primary tone indices rl_a and rl_b happens to have the largest amplitude among all other bins and the other tone bin happens to have highest amplitude among all other bins in a 2R Hz band surrounding the first located tone bin. The analysis of probability then proceeds in a similar manner to Algorithm I. First, the probability that the largest peak has an index equal to one of the primary modulation pair is P_{s1}. This will occur if one the primary tone bins happens to be the highest among all other bins.

$$
\begin{aligned}
P_{s1} = & \int_0^{\infty} \left[1 - \exp\left(\frac{-x^2}{2\sigma^2}\right)\right]^{N-4} \\
& \cdot \left[\int_0^{x} \frac{z}{\sigma^2} \exp\left(\frac{-(y^2+b^2)}{2\sigma^2}\right) I_0\left(\frac{by}{\sigma^2}\right) dz\right] \\
& \cdot \left[\int_0^{x} \frac{y}{\sigma^2} \exp\left(\frac{-(y^2+c^2)}{2\sigma^2}\right) I_0\left(\frac{cy}{\sigma^2}\right) dy\right]^2 \\
& \cdot \left[\frac{x}{\sigma^2} \exp\left(\frac{-(x^2+b^2)}{2\sigma^2}\right) I_0\left(\frac{bx}{\sigma^2}\right)\right] dx
\end{aligned}
\tag{3.27}
$$

The probability that the second primary tone is highest in a 2R Hz band centred on the first tone is P_{s2}.

$$P_{s2} = \int_0^{\infty} \left[1 - \exp\left(\frac{-x^2}{2\sigma^2}\right)\right]^{2R-3} \cdot \left[\int_0^{x} \frac{y}{\sigma^2} \exp\left(\frac{-(y^2+c^2)}{2\sigma^2}\right) I_0\left(\frac{cy}{\sigma^2}\right) dy\right] \cdot \left[\frac{x}{\sigma^2} \exp\left(\frac{-(x^2+b^2)}{2\sigma^2}\right) I_0\left(\frac{bx}{\sigma^2}\right)\right] dx \tag{3.28}$$

The probability of an FFT success is equal to the probability of index combinations (rl_a, rl_b) and (rl_b, rl_a). The probability of FFT failure is therefore equal to $P_{failure} = 1 - 2P_{s1}P_{s2}$. The simulation and analysis were implemented again using Matlab as shown in Appendix A.3. The accuracy of the analysis is demonstrated first for the probability of failure for an FFT of length 512 when compared to a simulation with 3000 trials as shown in Figure 3.15. The performance of algorithm I can then be compared with algorithm II as seen in the results in Figures 3.16 and 3.17 for N = 512. The results indicate that, as expected, Algorithm II has superior performance to Algorithm I at low signal to noise ratios. Both methods determine the carrier frequency with a low probability of error above 10 dB SNR in 1 MHz signal bandwidth for an FFT of length 512.

It should be noted that algorithm I provides a simpler implementation when used in the architecture shown in Figure 3.7 because it computes the frequency estimate in less steps than algorithm II. In fact Algorithm I could be implemented in the microcontroller by continuously updating the address pointers to the indices with the largest and second largest magnitudes every time a sample enters the microcontroller. This can be easily achieved as long the speed of the processor is faster than the delay between samples entering the microcontroller. Significantly, the estimation of the carrier frequency would take no more than a couple of μs after the last sample of the power spectrum has entered the processor. The DDS can then be set

to the correct frequency before the end of the pre-amble duration, t_w, to allow the Costas loop to demodulate the data transmission.

A slower, complicated but yet more accurate approach can implemented using algorithm II. Algorithm II requires that the second peak chosen must be the largest in a $2R$ Hz band centred on the index of the largest peak chosen from step one. This implies that the samples in this band must be ordered into descending numerical order. The sorting of these samples is not trivial and can only be performed after the last sample has entered the microcontroller. The number of operations for the simplest slowest sorting algorithms have $O(n^2)$ and the fastest algorithms have $O(n \log n)$. Whilst a discussion of these algorithms is beyond the scope of this work, lets consider a slow sorting algorithm operating on the samples in the band using a 16 MHz ATmega microcontroller. The number of samples in the band is dependent on the data rate and frequency resolution, if the data rate, R, is 1 MHz and f_{res} is 125 kHz, then the number of samples in the band is 16. Applying an algorithm complexity of $O(n^2)$ to sort these samples would take approximately 16 μs. This is in addition to the time of 8 μs taken to analyse the signal during the pre-amble duration, t_w, as well as any other latency caused by the settling times of the DDS and Costas loop used in the receiver design.

A much longer pre-amble duration must be accommodated when using algorithm II than in a system using algorithm I. The length of the pre-amble duration could be reduced by operating algorithm II using a faster processor which causes additional cost or reducing the frequency resolution which causes a higher probability that the algorithms would fail at low SNR. Moreover increasing the pre-amble length implies that the data burst will be much longer and thus the number of the RF-ID tags that can be accommodated reliably for a given mean update time will also be lower. These factors must be considered carefully before the implementation of algorithm II.

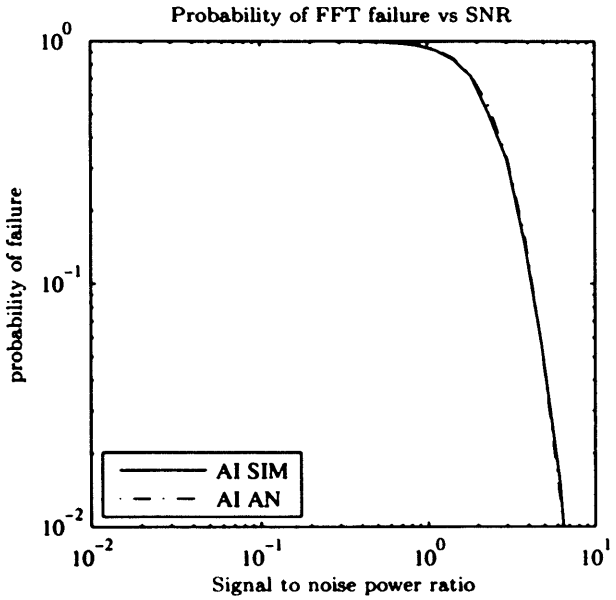


Figure 3.12: Simulated and analytic probability of failure using Algorithm I with an FFT size 512

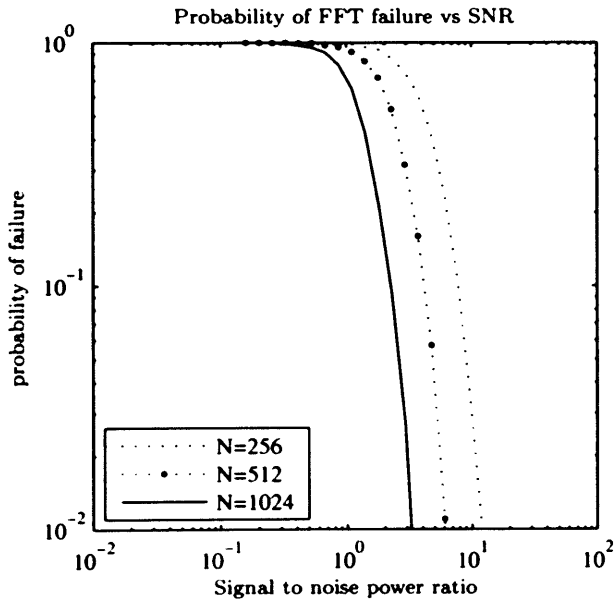


Figure 3.13: Simulated probability of failure using Algorithm I with FFT sizes of 256, 512 and 1024

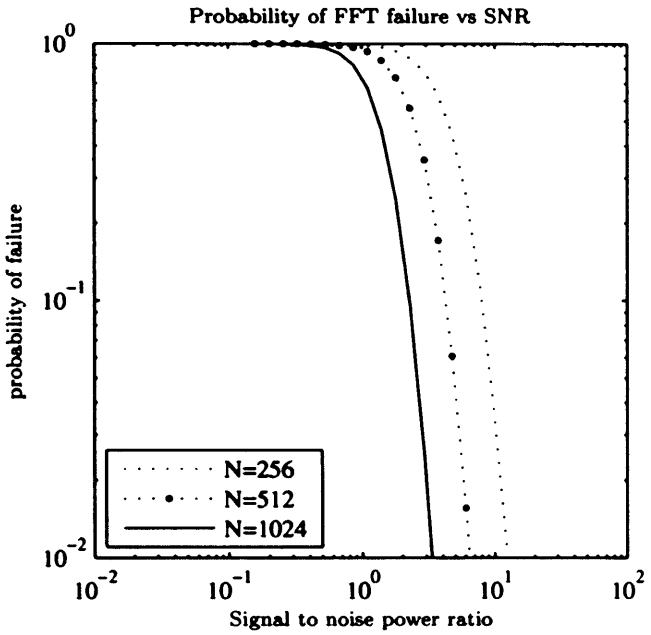


Figure 3.14: Analytic probability of failure using Algorithm I for FFT sizes of 256, 512 and 1024

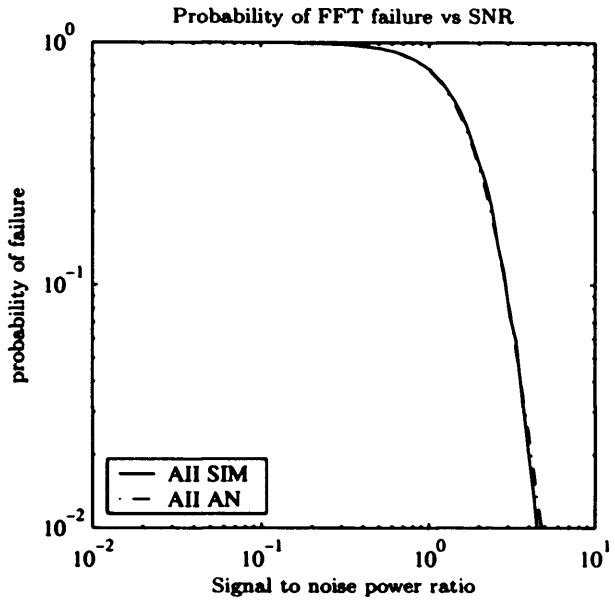


Figure 3.15: Simulated and analytic probability of failure using Algorithm II for an FFT size 512

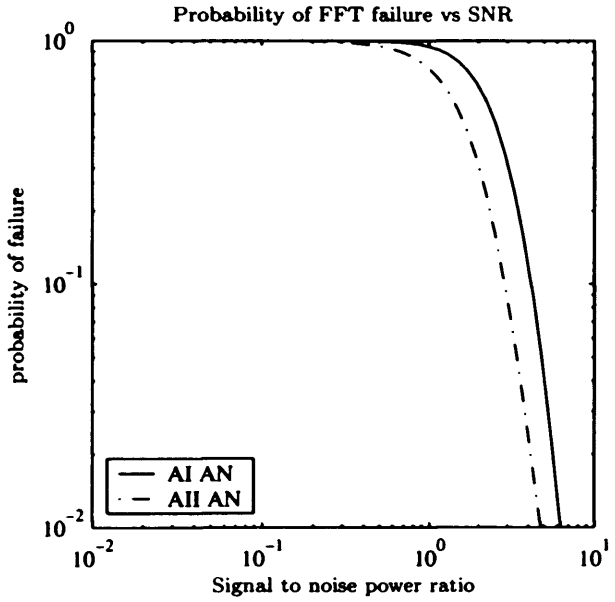


Figure 3.16: A comparison of the analytic probability of failure for Algorithms I and II for an FFT size of 512

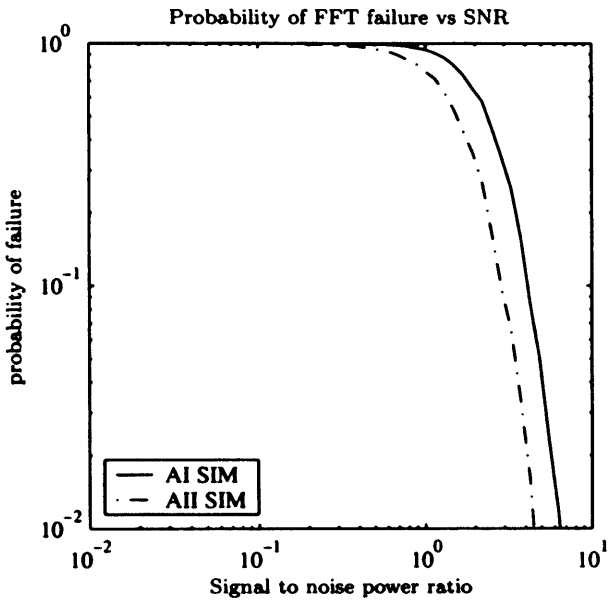


Figure 3.17: A comparison of the simulated probability of failure for Algorithms I and II for an FFT size of 512

3.3 Conclusions

The RF-ID readers can reliably provide demodulation and detection of the tag transmission at a ranges of 20 m or less even when the transmitted signal is severely affected by multipath and other degradations caused by the transmission channel. At ranges greater than 20 m the link margin is reduced and the system is more prone to error. Significantly, a large number of RF-ID tags, as many 1600 per cell, could be reliably identified with a mean update rate of one second using an asynchronous communication protocol. This protocol has one-way channel only and avoids the necessity of a return channel with acknowledgment messages, permitting, thus, the RF-ID tags to remain extremely simple. The protocol is also used in telemetry applications, although it is typically more useful for biomedical and ecological applications (where one would normally have an unidirectional data flow) than in engineering applications (where one usually has messages going both ways).

A novel RF-ID receiver architecture incorporating frequency diversity was also presented. The synchronisation technique used by the receiver has a carrier frequency estimator which is implemented here with SAW dispersive delay chirp filters as an alternative to a conventional FPGA or DSP implementation. The synchronisation technique contained in the receiver uses the spectral properties of a short pre-amble at the beginning of the received pulse to estimate the carrier frequency and then immediately directs a DDS-based LO to the appropriate frequency. The performance of the synchronisation technique on a BPSK modulated carrier signal has been examined for two peak search algorithms. Although the results indicate that Algorithm II has superior performance to Algorithm I at low signal to noise ratios, the implementation of Algorithm I leads to lower latency in the estimation of the carrier frequency because algorithm II uses an additional unavoidable step involving the sequencing of the data samples.

4. DEVELOPMENT OF THE ACTIVE RF-ID SYSTEM

The development of an active RF-ID system to be used in the Optag trial is demonstrated in this chapter. The design and development of the RF-ID tag is considered first followed by a description of the design and development of the reader which contains a number of receivers and a passive direction finding array. All aspects of the receiver design are covered including achieving maximum dynamic range with an AGC optimised for pulsed transmissions and maximising the direction finding accuracy of the antenna elements chosen for the array under non-ideal operating conditions of channel unbalance and additive Gaussian noise. The receiver chosen to be prototyped here does not contain any frequency diversity as considered in Chapter 3 because the demonstration of that technique is beyond the scope of this work. It is also apparent that the frequency interference is not anticipated within the transmission band in the locations used for the trial and moreover the additional complexity in the implementation of the RF-ID receiver caused by the incorporation of frequency diversity is not necessary in order to evaluate the effectiveness of the passive direction finding array.

We begin the evaluation of the tag design by presenting the options for the 5.8 GHz frequency source used by the transmitter. There are various methods for generating a signal at a frequency of 5.8 GHz including a free-running oscillator, a Dielectric Resonator Oscillator (DRO) and finally, a PLL frequency synthesiser based design. The first and second method both use an active device which can be either a high frequency FET or Bipolar transistor. They are widely used because they are simple to design, cheap, have a very low component count and can even built on standard FR4 PCB board. The DRO is more popular than the free running oscillator because it has better phase noise and minimum frequency drift as a result of the additional frequency stability created by the dielectric resonator. The design of the DRO for low phase noise and minimum frequency drift is critically dependent on the feedback mechanism. A frequently used mechanism is a common emitter transistor with a

series feedback configuration. This configuration has a resonant output frequency which is particularly sensitive to emitter inductance at 5.8 GHz which would create difficulties in reliably having all transmitters within the frequency capture range of the receiver. Besides the sensitivity of the resonant frequency to the emitter inductance, the resonant frequency is also susceptible to drift whenever the DRO circuit is disturbed by touching or slight movement. This is generally why DROs are housed in metallic boxes and are kept stationary. Unfortunately, methods one and two are not suitable for use in a portable application such as a tag. A more appropriate method uses a PLL (phase locked loop) to generate the desired RF carrier signal by using a reference oscillator. This method of generating the RF carrier signal has the advantage that it also has a high frequency stability and can not be easily disturbed, unlike the previous two methods.

4.1 PLL based transmitter

The general form of an integer PLL is shown in Figure 4.1. The loop is connected in a feedback arrangement in which the phase detector compares the input reference phase, ϕ_i, with the phase ϕ_o/N fed back from the VCO signal. The phase detector output voltage is dependent on the difference in phase between the input signals and adjusts the VCO until this phase difference is small. Whenever this point is reached the VCO phase is locked to the input reference phase, $\phi_o = N\phi_i$. Since frequency is the time derivative of phase, the circuit also behaves as a frequency multiplier and thus the output frequency may be stepped in multiples of the reference frequency, f_{ref}, by varying the division ratio, N.

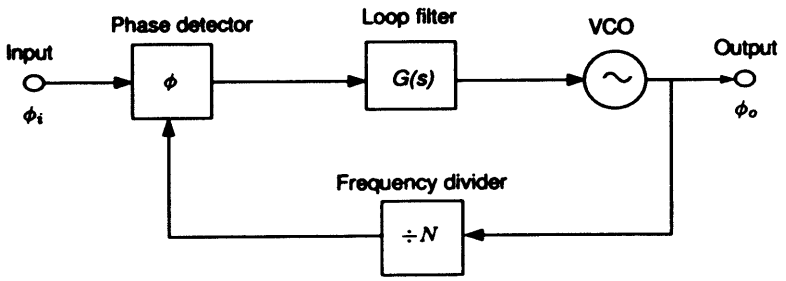


Figure 4.1: A block diagram of the basic PLL circuit

By operating the PLL at a high carrier frequency and with a low reference frequency the total division ratio will need to be very high. This presents a problem in maintaining low reference sidebands since any reference breakthrough is multiplied by N. Reference breakthrough occurs as result of the AC component of the phase detector output phase modulating the loop. It is undesirable and impairs the performance of the receiver. The problem can be addressed by a careful choice of loop filter. Let us for a moment examine a second order type II loop filter in a design with a reference frequency, f_{ref}, chosen as 15.625 kHz (i.e. for 125 kHz step size with a $\div$ 8 fixed prescaler). As a result the total division ratio, N, is found to be 371,200. As a rule of thumb the loop natural frequency for the design, f_n, is chosen as 1% of the reference frequency, f_{ref} and is set to 150 Hz. The final parameter in the design is a damping factor of $\xi = \frac{1}{\sqrt{2}}$. The reference breakthrough for such a design can be estimated using the peak phase deviation at the VCO output. The peak phase deviation at the VCO output is approximately equal to the phase deviation at the input of the phase detector, ϕ_{od}, multiplied by the total division ratio, N, multiplied by the closed loop response of a second order type II for PM modulation [22],

$$\Delta\phi \approx \Delta\phi_{od} \times N \times \left(\frac{2\xi\omega_n}{\omega_{ref}}\right) \tag{4.1}$$

For a realistic design, $\phi_{od} \approx 1°$, which means that the peak phase deviation at the output of the VCO is $\Delta\phi = 4965° \equiv$ 100 rads. A modulation index of 100. The level of reference sidebands would be very high and this would be a disasterous design. An alternative approach to help reduce the reference sidebands is a third order type II filter with additional low pass filtering of $20f_n$ on the op amp output. This time the loop natural frequency, f_n was lowered to a more cautious value of 50 Hz with reference frequency, f_{ref}, remaining as 15.625 kHz, and total division ratio, N, also remaining as 371,200. The peak phase deviation at the VCO output can be calculated in similar way to previously only this time using the closed loop response of a third order type II for PM modulation [22] and the closed loop response of a

low pass filter,

$$\Delta\phi = \Delta\phi_{od} \times N \times \frac{\omega_n^2 (\tan\phi + \sec\phi)}{\omega_{ref}^2} \times \frac{1}{15.625} \tag{4.2}$$

Again for a realistic design, $\phi_{od} \approx 1°$, which means that the peak phase deviation at the output of the VCO is now $\Delta\phi = 0.89° \equiv 0.0155$ rads, i.e a modulation index of 0.0155. Converting to sideband levels, using $20\log_{10}\left(\frac{\Delta\phi}{2}\right)$ the first harmonic sideband level is -42.2 dBc which is significantly better than the previous design but can be improved further. It would help if f_{ref} is increased to 31.25 kHz. This would provide a further x 8 increase in filter attenuation and a two fold reduction in N. In total this provides a 24 dB reduction in reference sidebands to circa -66 dBc at the first harmonic.

The combined transfer function of the 3rd type II loop filter and low pass filter shown in Figure 4.2(a) is

$$G(s) = \frac{1}{s\tau_1} \cdot \frac{1 + s\tau_3}{1 + s\tau_2} \cdot \frac{1}{1 + s\tau_4} \tag{4.3}$$

where $\tau_1 = R_1C_1$, $\tau_2 = R_2C_2$, $\tau_3 = R_2(C_1 + C_2)$, and $\tau_4 = R_3C_3$. The time constants are dependent on the required design. The design values can be calculated from the following equations [22].

$$R_1C_1 = \frac{K_pK_v}{N\omega_n} (\tan\phi + \sec\phi) \tag{4.4}$$

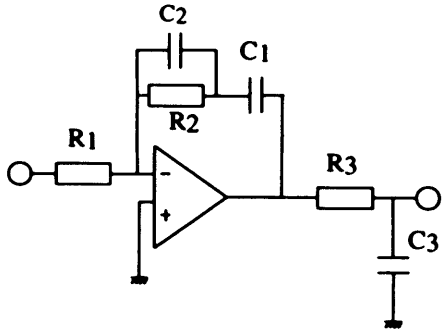
$$R_2C_2 = \frac{1}{\omega_n (\tan\phi + \sec\phi)} \tag{4.5}$$

$$R_2C_1 = \frac{2\tan\phi}{\omega_n} \tag{4.6}$$

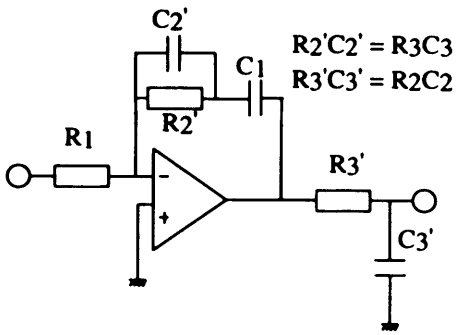
$$R_3C_3 = \frac{1}{20\omega_n} \tag{4.7}$$

where K_v is the VCO gain of $2\pi \times 100$ MHzV^{-1}, K_p is the phase detector gain of $\frac{3}{2\pi}$ Vrad^{-1}, ω_n is the loop natural frequency of 314 rads^{-1}, N is the total division ratio of 185,600 and ϕ is the phase margin of 60°. Using these equations it was

found that $\tau_1 = 67$ ms, $\tau_2 = 853$ μs, $\tau_3 = 11$ ms and $\tau_4 = R_3C_3 = 0.159$ ms. Since $\tau_2 < \tau_4$ it would make sense to swap the time constants as illustrated in Figure 4.2(b) for some additional filtering of the op-amp noise whilst retaining the same loop characteristic.



(a) Loop filter with transfer function, $G(s)$



(b) Loop filter with transfer function, $G(s)$ with τ_2 and τ_4 swapped

Figure 4.2: The third order type II loop filter used in the PLL

The design for the single channel PLL frequency source is shown as Figure 4.3. The 50 Hz loop design ensures that the loop was stable and would retain lock but at the cost of regaining lock more slowly if the loop was disturbed. The PLL was successfully tested and found to lock correctly and stably over the full range of frequencies. The phase noise was then tested and is defined as the ratio of the carrier power to noise power found in a 1 Hz bandwidth at a defined frequency offset. Expressed in dBcHz^{-1}, the inband (or close-in) phase noise is dominated mainly by the VCO noise contribution with a small contribution from the synthesiser. After viewing the frequency output from the prototype PLL frequency source on a spectrum analyser it was found that the phase noise was -65 dBc @ 500 kHz offset in a 300 kHz resolution bandwidth. Taking $10\log_{10}$ of the resolution bandwidth we find that the phase noise is -120 dBcHz^{-1} @ 500 kHz offset. As expected from this design, the reference sideband level for the first harmonic was very low and was measured to be approximately -65 dBc.

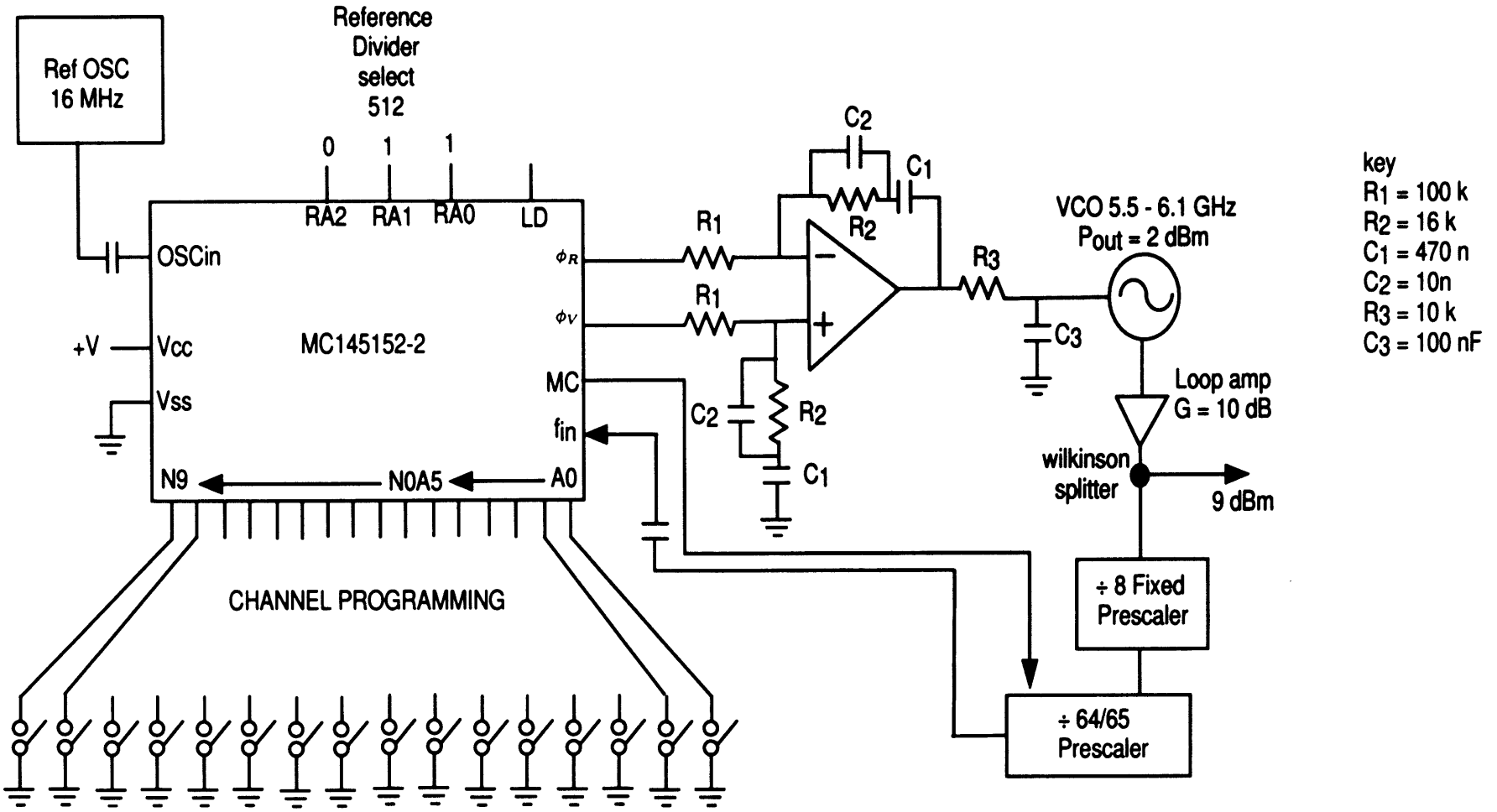


Figure 4.3: A schematic of the prototype frequency synthesiser

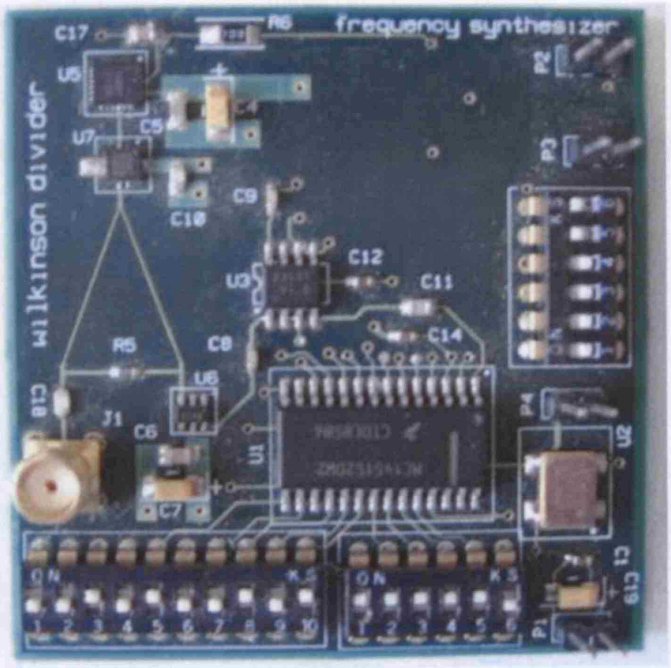


Figure 4.4: A photo of the prototype frequency synthesiser

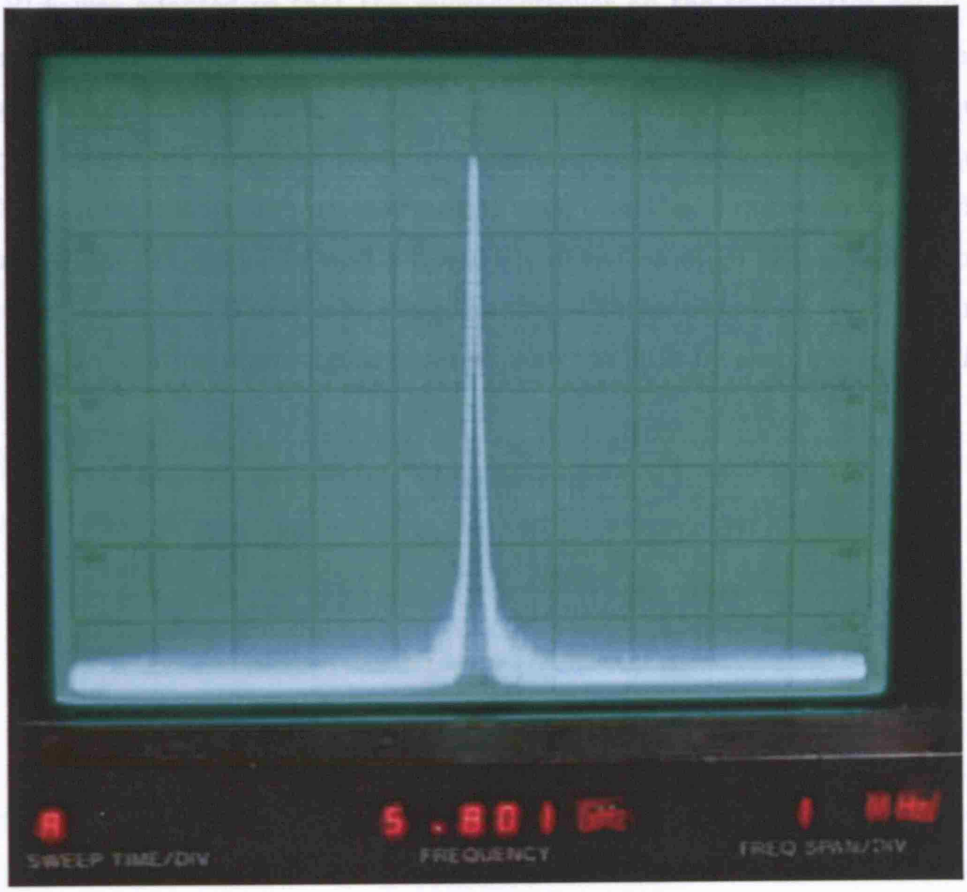


Figure 4.5: A snapshot of the output signal spectrum from the prototype frequency synthesiser

The PLL was adapted so that the microcontroller on the transmitter could modulate the VCO tuning input and so frequency shift key the data onto the carrier signal. This could be achieved using a voltage divider since the maximum frequency component in a data rate of 100 kbps is much greater the loop natural frequency of 50 Hz. The voltage divider is arranged such that the loop dynamics previously calculated are not disturbed and a frequency deviation of 50 kHz is present on the output of the VCO. The modulating interface can be described using Figure 4.6 with the modulating input signal isolated from the PLL by ensuring $R_3 >> R_2$.

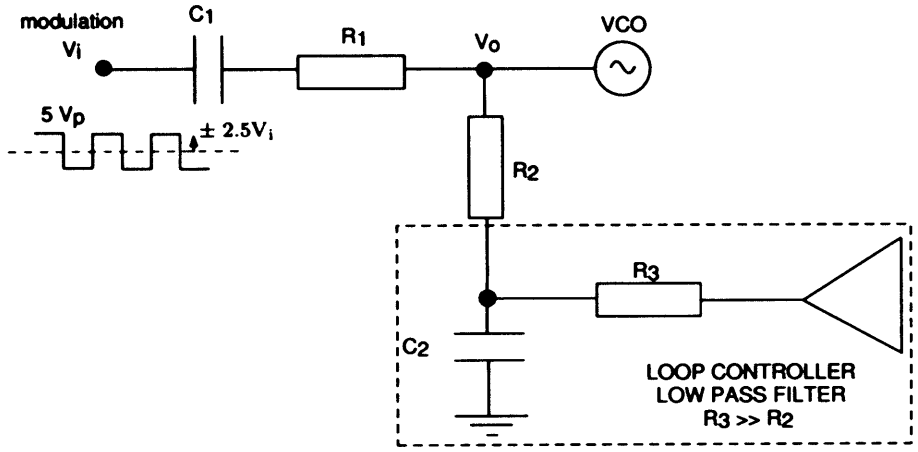


Figure 4.6: A schematic illustrating how the PLL loop is adapted for FSK modulation

The frequency response of the voltage on tuning input of the VCO, V_o, to the modulation input voltage, V_i, from the microcontroller can be described as

$$\begin{aligned}
\frac{V_o}{V_i} &= \frac{R_2 + \dfrac{1}{sC_2}}{R_1 + R_2 + \dfrac{1}{sC_1} + \dfrac{1}{sC_2}} \times \frac{sC_2}{sC_1C_2} \times C_1 \\
&\equiv \frac{C_1\left(1 + sC_2R_2\right)}{C_1 + C_2 + sC_1C_2\left(R_1 + R_2\right)} \\
&\equiv \frac{C_1}{C_1 + C_2} \bullet \frac{1 + sC_2R_2}{1 + \dfrac{sC_1C_2\left(R_1 + R_2\right)}{C_1 + C_2}}
\end{aligned} \tag{4.8}$$

The magnitude of this function and the location of pole and zero are now shown

in Figure 4.7 for the components resistors R_1,R_2 and capacitors C_1,C_2. At low frequency the capacitors act as DC blocks and the transfer function is equivalent to the capacitor ratio of

$$\frac{C_1}{C_1 + C_2}$$

On the other hand at higher frequencies the capacitors act as short circuits and the output is equivalent to the resistor ratio of

$$\frac{R_2}{R_1 + R_2}$$

As a result of unequal resistor and capacitor ratios, the break points correspond to a zero at

$$\frac{1}{R_2 C_2}$$

and a pole at

$$\frac{C_1 + C_2}{C_1 C_2 (R_1 + R_2)}$$

are at different frequencies. In this example the pole break point is lower in frequency than the zero break point. The variation in the amplitude of the frequency response could cause the peak frequency deviation on the output of the VCO to vary with the instantanous frequency of the data transmitted by the tag, particularly if there were long strings of 0s and 1s present in the data. To ensure that the frequency response is flat over all input frequencies the capacitor and resistor ratios were made equal.

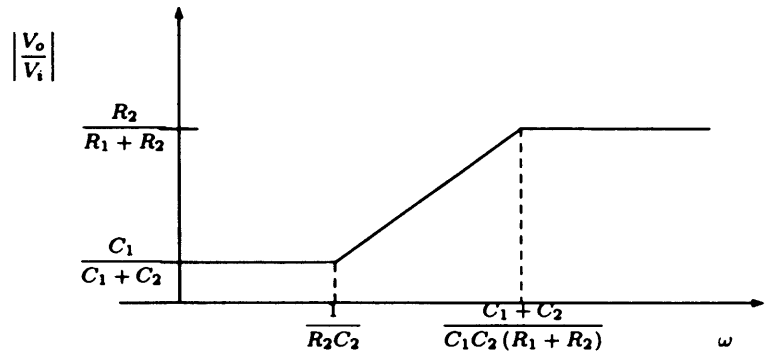


Figure 4.7: Frequency response of the voltage divider used for FSK modulation of the PLL

The condition which sets the capacitor and resistor ratios equal for a flat frequency response can be described in the following way

$$\begin{aligned} \frac{R_2}{R_1 + R_2} &= \frac{C_1}{C_1 + C_2} \\ R_2C_1 + R_2C_2 &= R_1C_1 + R_2C_1 \\ R_1C_1 &= R_2C_2 \end{aligned}$$

Before we can satisfy this condition using the appropriate values for capacitor C2 and resistor R2 the divider ratio must first be calculated from

$$\frac{R2}{R1} = \frac{\Delta\phi}{k_v \times V_{i\,max}}$$

For a desired peak frequency deviation, $\Delta\phi$, of 50 kHz, a VCO sensitivity, K_v, of 100 MHz / volt, and a maximum AC coupled input voltage, V_{imax}, of 2.5 volts, the divider ratio is 1/5000. We can now set the resistor values so that $R_2 << R_3$ by choosing $R_1 = 1$ M Ω and $R_2 = 200\ \Omega$. Since capacitor C_2 is determined by the low pass filter time constant in the PLL as 100 nF the only remaining value to set is capacitor C_1. As described earlier this must be less in value than capacitor C_2 and ideally smaller by a factor of 5000, in the design it is chosen as 20 pF for convenience. As a result of these design values the frequency response is approximately flat across all modulating frequences in the range from DC to 50 KHz.

These design values are now used to modify the original PLL as shown in Figure 4.8. The modified PLL frequency source was then incorporated into the tag as shown in the completed prototype board in Figure 4.11. While the tag is not transmitting a data burst the PLL maintains a mean centre frequency by modulating the VCO with alternate ones and zeros. The mean centre frequency was matched to a voltage threshold in the FSK discriminator in the receiver that determines the logic state of the data. If the mean centre frequency was not maintained by the PLL, a step change in frequency applied to the VCO tuning input at the beginning of a databurst would take 3.41 ms to translate to a settled output. This causes the mean centre frequency to vary over the 1.28 ms length of the data burst and this variation would adversely affect the performance of the FSK discriminator. Figures 4.9 and 4.10 show the output power on a spectrum analyser for a CW signal and FSK modulated carrier signal from the prototype transmitter pcb board. The tag uses FSK modulation and has an output power of 7 dBm at 5.8 GHz with a phase noise of -120 $dBcHz^{-1}$ @ 500 kHz offset. The data rate is set at 100 kbit/s and consequently there are spectral components at 50 kHz spacing as shown in Figure 4.10. For this tag design the peak frequency deviation is also set at 50 kHz which is equivalent to a modulation index of one.

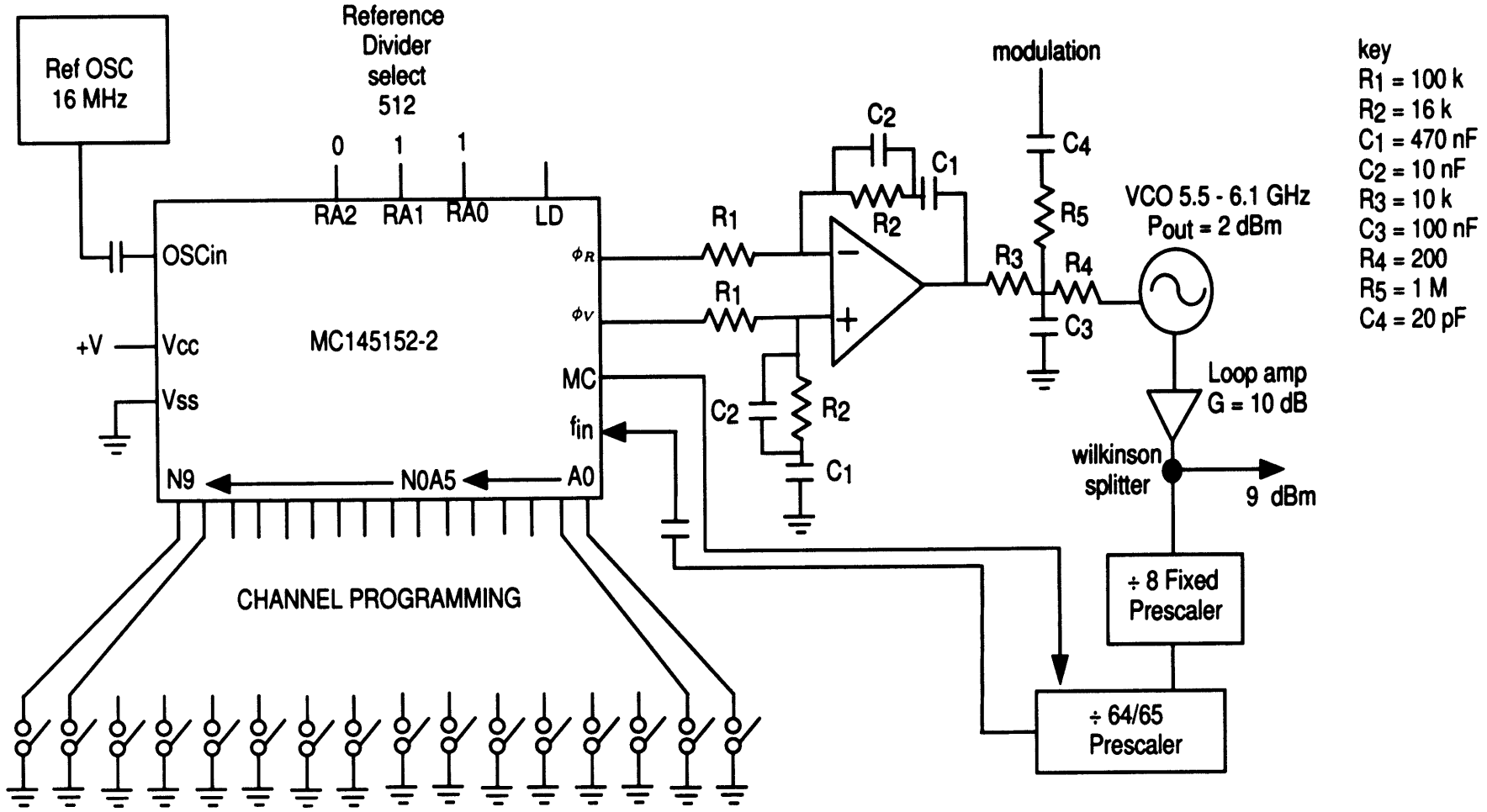


Figure 4.8: A schematic of completed frequency synthesiser with modification for FSK modulation

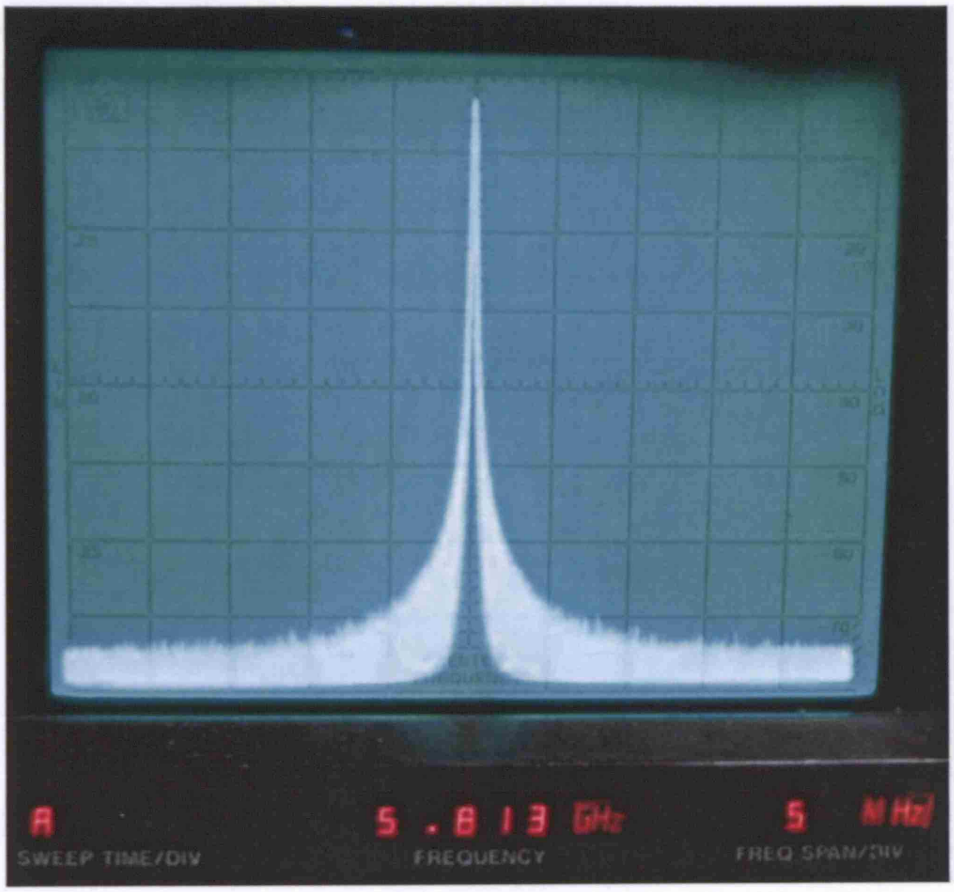


Figure 4.9: The output power spectrum from the completed tag for a CW signal

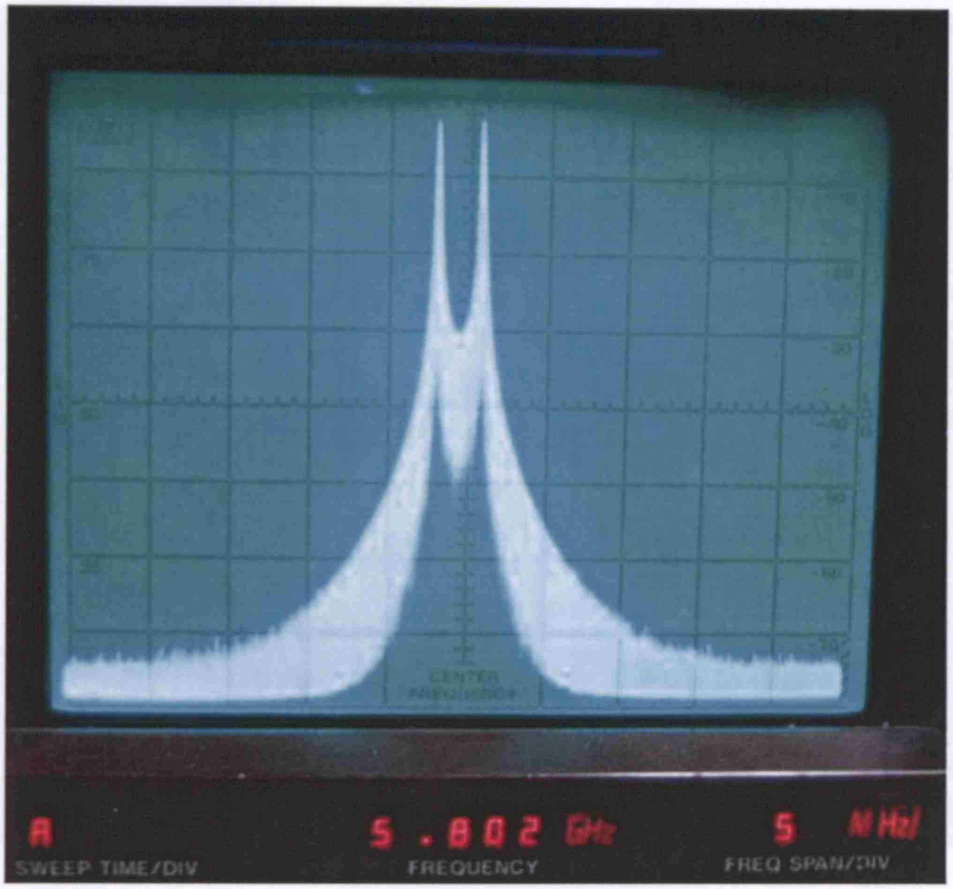


Figure 4.10: The output power spectrum from the completed tag for a FSK modulated carrier signal with modulation parameters – f_m of 50 kHz (100 kb/s data stream consisting of alternative zeros and ones) and Δf of 50 kHz

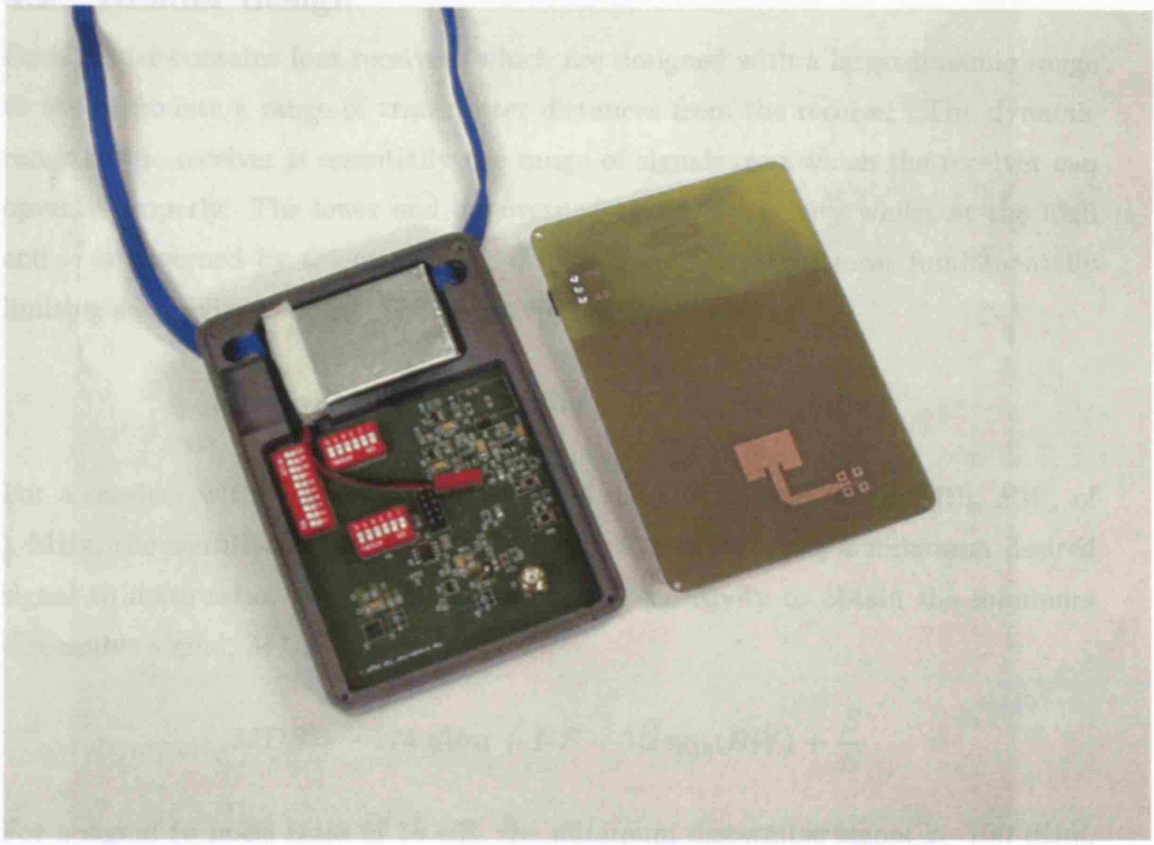


Figure 4.11: The completed tag with covering, battery and antenna lid

4.2 Reader design

Each reader contains four receivers which are designed with a large dynamic range to accommodate a range of transmitter distances from the receiver. The dynamic range of the receiver is essentially the range of signals over which the receiver can operate properly. The lower end is governed by the sensitivity whilst at the high end it is governed by the strong signal handling. The factor most fundamentally limiting sensitivity is noise. The sensitivity can be stated as

$$\text{Sensitivity } (S) = -174 \text{ dBm} + NF + 10\log_{10}(BW)$$

For a receiver with a noise figure, NF, of 4 dB and a signal bandwidth, BW, of 1 MHz, the sensitivity is -110 dBm. Usually the receiver has a minimum desired signal to noise ratio, this can be added to the sensitivity to obtain the minimum discernible signal, MDS as

$$MDS = -174 \text{ dBm} + NF + 10\log_{10}(BW) + \frac{S}{N}$$

For a signal to noise ratio of 10 dB, the minimum discernible signal is -100 dBm. Now that we know the MDS we can plot the expected range of the receiver using the link budget. The link budget can be approximated by

$$P_r \text{ (dBm)} = P_t \text{ (dBm)} - FSL(\text{dB})$$

$$FSL = 20\log 10\left(\frac{4\pi d}{\lambda}\right)$$

Assuming an EIRP, P_t, of 10 dBm, a wavelength, λ, of 5.17 cm the received power, P_r, can be plotted versus distance, d ,as shown in Figure 4.12.

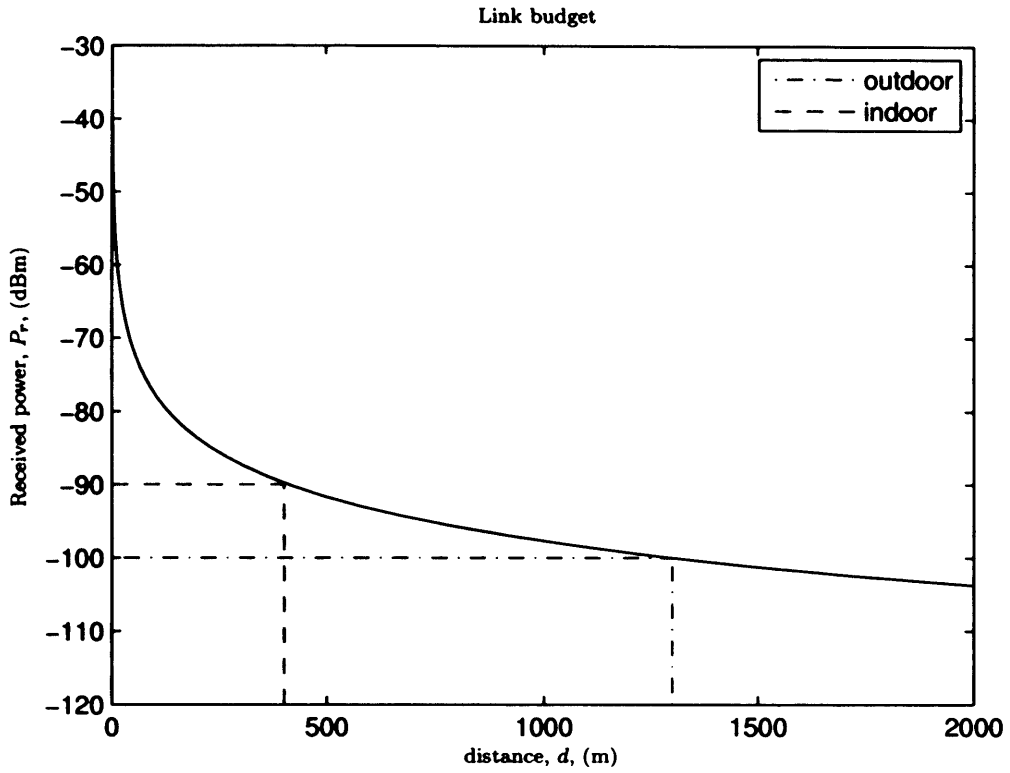


Figure 4.12: The link budget assessed indoors and outdoors using the received power versus range for each receiver

The receiver range is just over 1300 m in an outdoor environment with ideal propagating conditions for an MDS of -100 dBm whilst the receiver range in an indoor environment is likely to be much less, for instance if a 10 dB standard deviation is applied to the path loss formula then a range of just over 400 m can be expected. However it is likely the propagation conditions in an indoor environment will be more severe and the range will be degraded to a greater extent. Having defined the sensitivity and range we now determine the maximum signal to be received based on the likely minimum distance of the transmitter from the receiver. Assuming a near distance of 1 m, the received power is found to be -33 dBm. The dynamic range required from the receiver can now be calculated from Pmax (-33 dBm) - Pmin (-100dBm) to be 67 dB.

The received signal by the antenna is first passed to the RF front end. The schematic and picture of the RF front end PCB is shown in Figure 4.2 and 4.18. The RF front end board down converts the received narrow-band FSK modulated carrier from 5.8 GHz to a differential FSK 12 MHz signal using an image rejection mixer with a low injection LO frequency signal. Before entering the mixer the carrier is amplified by 3 LNAs with a total gain of approximately 35 dB. The noise figure of this cascaded chain of amplifiers is set primarily by the first amplification stage. Consequently the first amplifier was set for a low noise figure whilst the next stages of amplification were set for maximum gain. This led to an overall front end noise figure of 4 dB. The LO signal from a synthesiser board (4 channel PLL) was also amplified from -10 dBm to a drive level greater than 10 dBm on the input to the mixer using two stages of amplification which were both set to maximum gain. Provided the mixer was sufficiently driven with at least 10 dBm from the LO, a conversion loss of only 8 dB was expected. The RF front end board was also tested and the RF to IF gain was measured to be approximately 27 dB in line with expectations. In addition the RF front end was found to exhibit a linear change in IF signal level to changes in the RF signal level. This can be seen in Figures 4.14 and 4.15. The 12 MHz IF differential output was examined for an LO drive level of -10 dBm at 5.802 GHz and RF input levels of -30 dBm and -50 dBm at 5.814 GHz. The board was then examined for image isolation and was found to have a measured image rejection performance of about 30 dB.

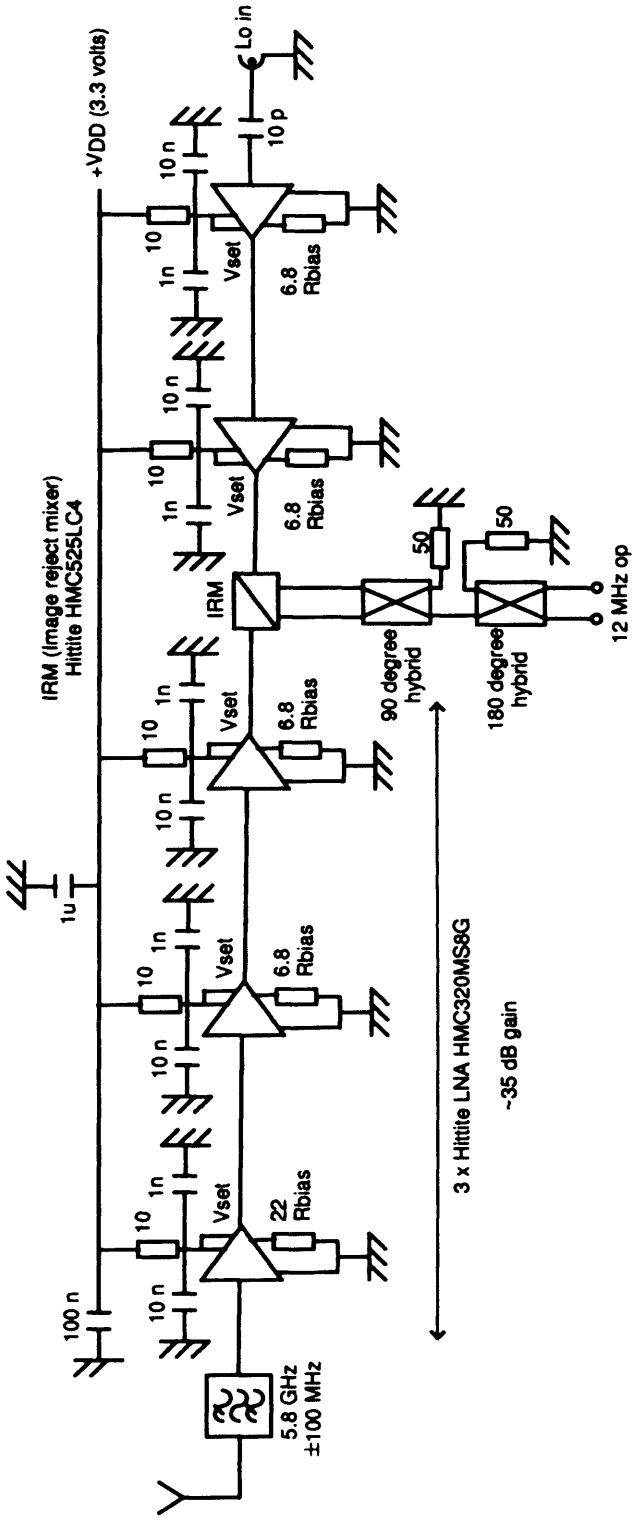


Figure 4.13: A schematic of the RF front end used in each RF-ID receiver

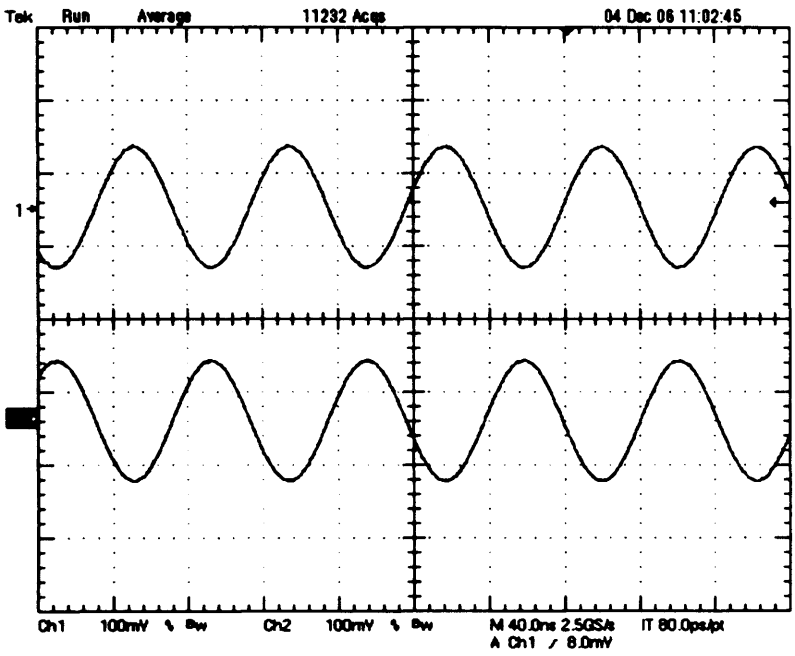


Figure 4.14: A plot of the measured IF output signal from the RF front end for an RF input power level -30 dBm

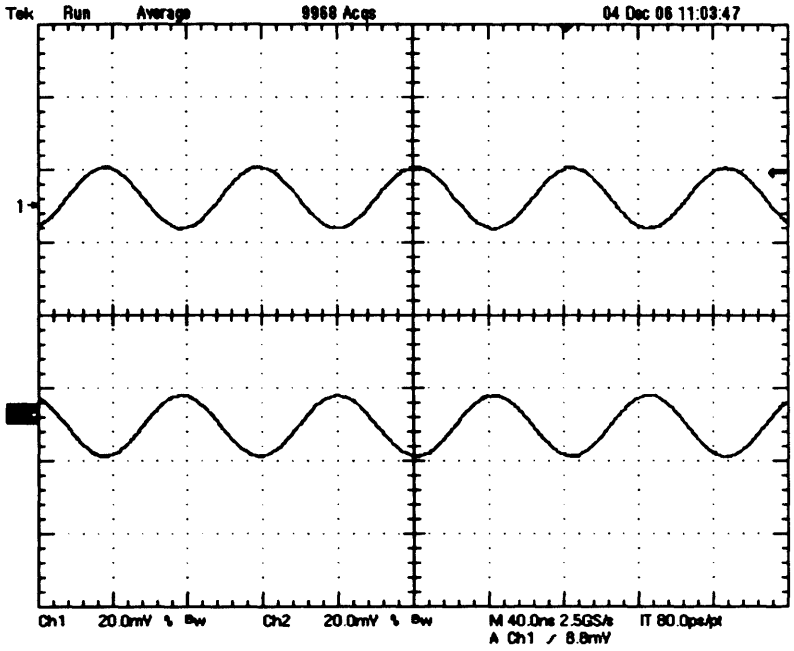


Figure 4.15: A plot of the measured IF output signal from the RF front end for an RF input power level -50 dBm

The 12 MHz differential IF signal is then passed to an AGC amplifier that consists of an automatic gain control circuit with 80 dB of dynamic range with a linear-in-dB received signal strength indicator (RSSI). The AGC is shown in Figure 4.20 and operates in differential mode to mitigate potential cross-talk between the four channels, which are packed in close proximity in the reader enclosure. AGCs normally operate on continuous signals using rather long loop time constants of around 1 ms. In contrast the tag signals here are pulsed and the AGC is required to lock onto the tag burst before the end of the pre-amble. For a tag pre-amble of 16 bits and a symbol time 10 μs, the loop time constant must be significantly less than 160 μs. Such a short time constant makes the AGC loop more susceptible to instability since it may over react to transient behaviour. The transient behaviour could be caused by signal feedback between stages and this meant that the gain stages of the AGC and the filtering in the design had to be precisely set to ensure that a stable RSSI (received signal strength estimate)was sampled by the AGC. The behaviour of the AGC loop is described in further detail in the the block diagram shown in Figure 4.16. In the design considered here the diode voltage response has a square-law characteristic where the instantaneous output voltage is K_d times the square of the input voltage and, equally, the mean value of the detector output voltage is equal to K_d times the square of the rms value of the sinusoidal input voltage. K_a is the VCA gain characteristic in dB/V and A is the gain at zero control voltage. In this analysis K_o implies the rms value and K_{det} implies the mean value.

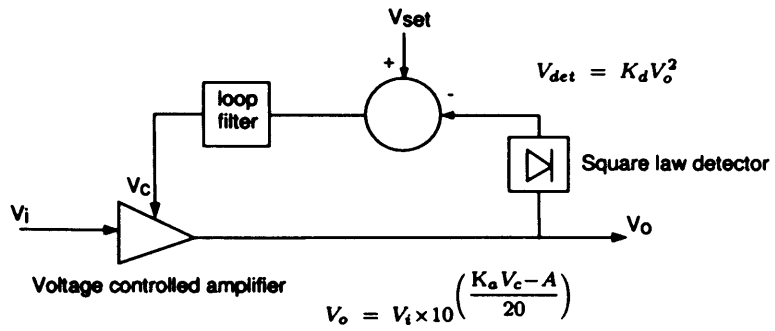


Figure 4.16: The architecture of the AGC loop used in the RF-ID receiver

In order to design the loop filter it is necessary to make a linearising approximation by differentiating the VCA and the square law detector characteristics and considering AGC operation under small signal conditions.

$$\begin{aligned} \frac{\delta V_o}{\delta V_c} &= \frac{K_a In(10) V_o}{20} \\ \frac{\delta V_{det}}{\delta V_o} &= 2K_d V_o \end{aligned}$$

These transfer functions can now be applied to the signal flow graph description of the circuit, in Figure 4.17, where it is assumed that the loop filter is based on a simple integrator given by the transfer function $-\frac{1}{s\tau}$, where $\tau = RC$ is the integrator time constant. The resulting loop response is given by

$$\frac{V_o(s)}{V_{set}(s)} = \frac{\frac{K_a In(10) V_o}{20\tau}}{s + \frac{K_d K_a In(10) V_o^2}{10\tau}} \tag{4.9}$$

which describes a first order low pass response with a 3 dB cut-off frequency of

$$\omega_o = \frac{K_a K_d In(10) V_o^2}{10\tau} \tag{4.10}$$

The loop filter is designed for operation when the loop is settled, at which point the

output voltage is $V_o = \sqrt{V_{set}/K_d}$. It should be noted the set voltage defines the AGC fixed output voltage when it is locked. This is chosen to be a single ended power level of 0 dBm, equivalently 0.223 volts rms or 0.632 volts pp.

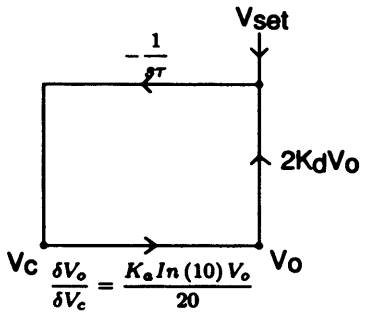


Figure 4.17: A signal flow representation of the AGC loop

The single ended output from the AGC which is set to limit at 0 dBm is then passed to a limiting amplifier of 10 dB gain. The output of the limiter is a 10 dBm signal with at least a 10 dB signal to noise ratio that is passed to an FSK discriminator integrated on the same board as the AGC. The integrated FM discriminator works using two balanced AM demodulators, one tuned with a high pass response and the other a low pass response around the 12 MHz centre frequency. The differential responses are then combined and compared to a voltage threshold using a comparator to determine the data. The output from the board is a digital bitstream at 100 kb/s which is then passed to a computer via a decoder shown in the Figure 4.21 and an ethernet interface (not shown). The RSSI function is provided by sampling the analogue AGC control voltage, which gives a linear-in-dB measure of the signal strength. Before the voltages could be converted to signal strengths, each receiver was calibrated to identify any voltage offsets from the designed linear response of 40 dB per volt slope and a calibrated reference voltage of 0.65 volts for -30 dBm input signal power. The calibration also ensured that any voltage offsets present between channels were accounted for whenever the AOA (Angle of arrival) was computed from the difference in voltage between the strongest two adjacent channels.

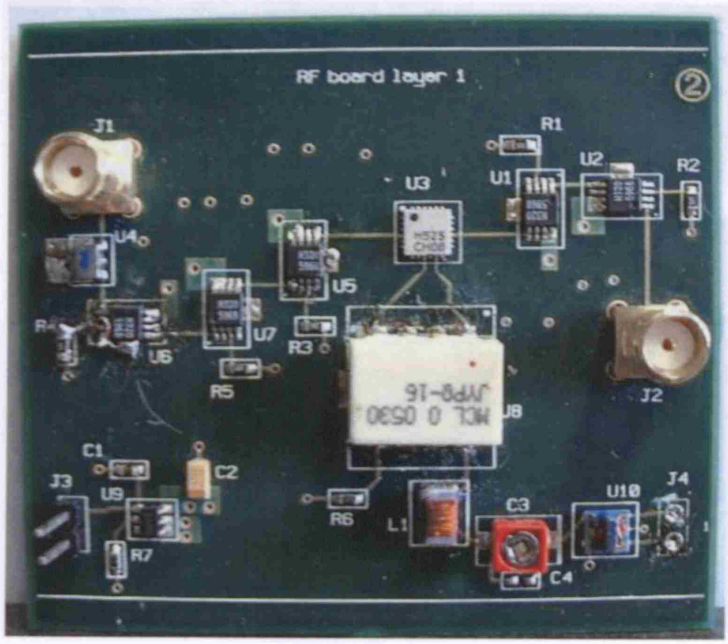


Figure 4.18: The RF front end PCB used in the tag reader

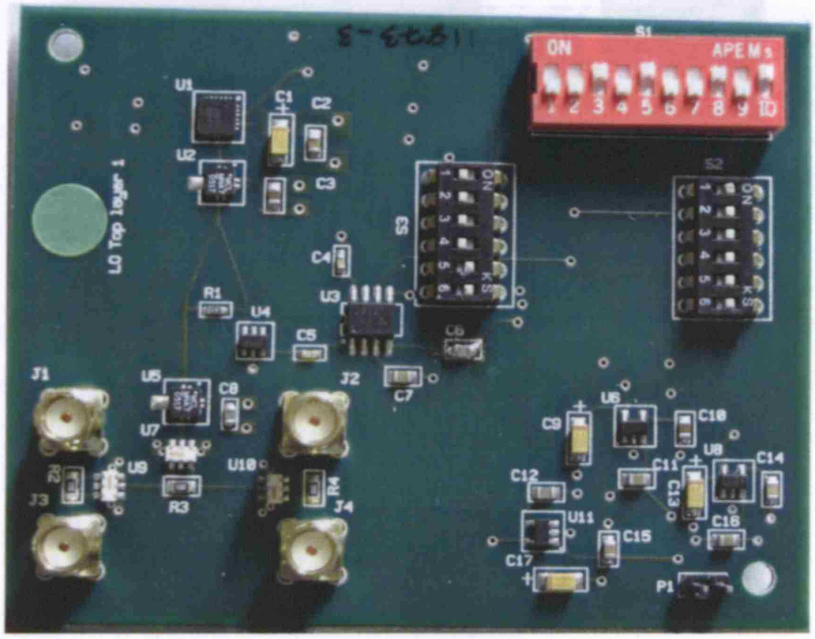


Figure 4.19: The synthesiser PCB used in the tag reader

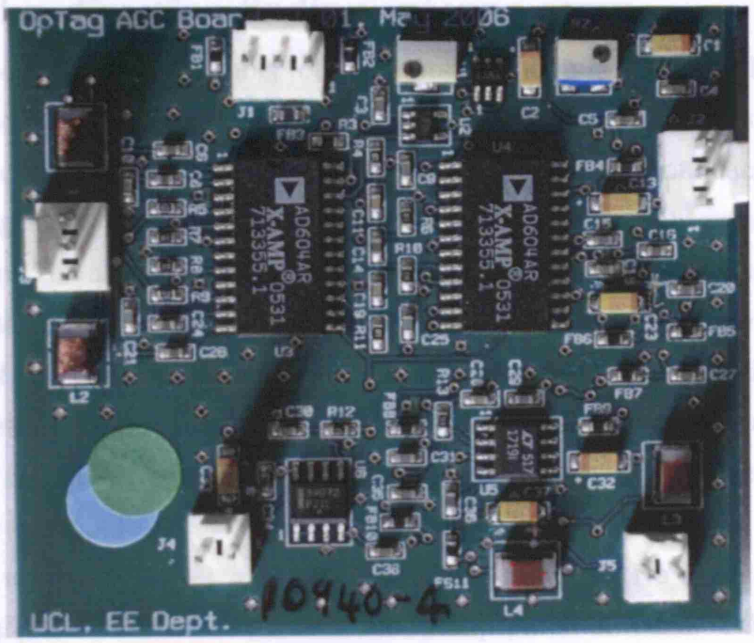


Figure 4.20: The AGC/FSK discriminator PCB used in the tag reader

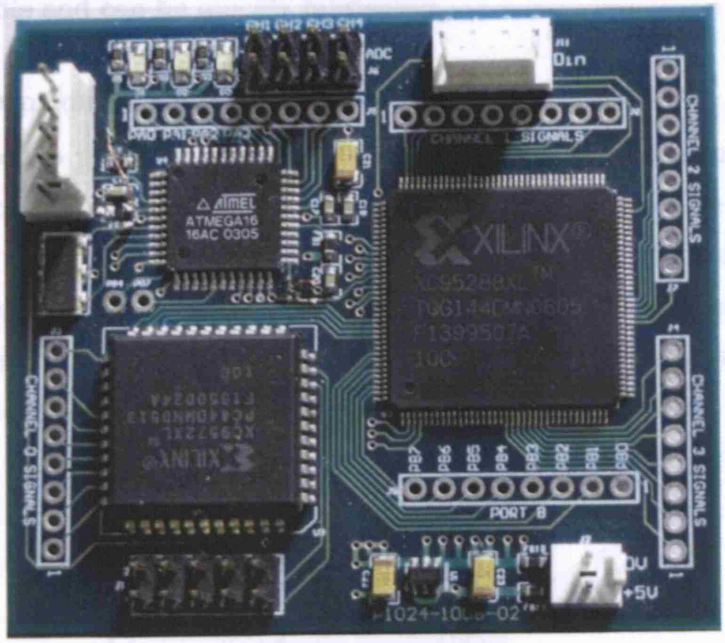


Figure 4.21: The decoder PCB used in the tag reader

4.3 Passive direction finding antenna array

Each reader consists of a number of receivers containing antenna elements with directional radiation patterns arranged in an array with a 360° field of view. The antenna elements are squinted at regular intervals in azimuth from each other. The regular arrangement allows the power ratio between adjacent antenna elements to uniquely define the angle of arrival in azimuth. The power ratio is less accurate in estimating the angle of arrival whenever it varies in azimuth and elevation. This could occur if the transmitter and receiver are located at different heights or have different pointing angles. As a consequence, ideally the radiation pattern would have a rapid roll off in azimuth and a slow roll off in elevation to compensate for the elevation error. The horn antenna has this characteristic when finely tuned but it major weakness is that it is not compact. An alternative more cost effective approach is a planar antenna such as a microstrip patch antenna. While this type of antenna has a slightly faster roll off in elevation than a horn antenna, at present implementing the radiation pattern using a patch is more convenient since it is light, has a low profile and can be quickly fabricated.

4.3.1 Description of the antenna element

A microstrip patch antenna consists of a planar substrate material with a radiating patch on one side and a ground plane on the other. The patch usually has the shape of a rectangle but it can also take on the dimensions of a circle, and be constructed on curved surfaces such as a cylinder. The radiation pattern of a simple edge fed rectangular microstrip patch shown in Figure 4.22 can be modeled by treating the fringing fields along the edges as radiating apertures [27].

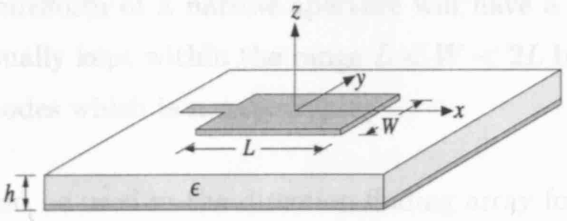


Figure 4.22: An illustration of a edge fed Microtrip patch

In that analysis the normalised gain is given by

$$g(\theta,\phi) = \frac{|\mathbf{E}(\theta,\phi)|^2}{|\mathbf{E}|^2_{max}} = \left(\cos^2\theta\sin^2\phi + \cos^2\phi\right)|\mathbf{F}(\theta,\phi)|^2 \tag{4.11}$$

where $\mathsf{F}(\theta,\phi)$ is the radiation vector. The E-plane and H-plane gains are obtained by setting $\phi = 0$ and $\phi = 90°$.

$$\begin{aligned} g_{\mathbf{E}}(\theta) &= \frac{|\mathbf{E}_\theta|^2}{|\mathbf{E}_\theta|^2_{max}} = |\cos(\pi v_x)|^2,\ v_x = \frac{L}{\lambda}\sin\theta \\ g_H(\theta) &= \frac{|\mathbf{E}_\phi|^2}{|\mathbf{E}_\phi|^2_{max}} = \left|\cos\theta\frac{\sin(\pi v_y)}{\pi v_y}\right|^2,\ v_x = \frac{W}{\lambda}\sin\theta \end{aligned} \tag{4.12}$$

The resonant length, L, of the patch is a function of the wavelength, λ and the dielectric constant, ε_r and is approximately given by

$$L = 0.5\frac{\lambda}{\sqrt{\varepsilon_r}} \tag{4.13}$$

The radiation pattern for the E-plane and H-plane are now plotted in Figure 4.23 for a patch with a FR4 substrate of bulk dielectric constant, ε_r, of 4.4 and width, W, equal to one and half times the length, L. The H-plane has a narrow radiation pattern with a beamwidth of 78° whilst the E-plane has a broad radiation pattern with a beamwidth of 180°. Compared with the normalised $\cos^2(\phi)$ and $\cos(\phi)$ radiation patterns we see that the H-plane has a roll off similar to the $\cos(\phi)$ radiation pattern. In fact the beamwidth H-plane can be adjusted by varying the width, W, since the Fourier transform of a wider aperture will have a narrower beamwidth while the Fourier transform of a narrow aperture will have a broader beamwidth. The width, W, is usually kept within the range $L < W < 2L$ because larger widths may cause higher modes which is not desirable.

The H-plane should be used in the direction finding array for reasonable DF resolution and greater mitigation of multipath signals because it has much narrow beamwidth than the E-plane. Figure 4.24 shows the H-plane power ratio for a pair

of adjacent patch antennas squinted by 90° in azimuth and having zero tilt in the elevation. Also plotted is the H-plane power ratio recalculated for elevation angles of -5° and -10°. It is clear that there will be uncertainties in the estimation of AOA from the measured power ratio in dB for a signal received at an unknown elevation angle because without prior knowledge of the elevation angle of the received signal the power ratio function evaluated at zero elevation must be used to estimate the AOA. Specifically the error is only significant outside the ±10 dB range and is no more than 10° over a range ±16°, in AOA.

Ideally the logarithmic power ratio slope in dB between adjacent elements should vary in a linear fashion with AOA and be invariant with elevation angle. Direction finding arrays which have this property usually contain antenna elements with exponential radiation patterns and can be operated over a wide frequency band. This property is also are very desirable for direction finding applications because it allows a simple digital look up table for the assessment of angles of arrival. As seen in Figure 4.24 the azimuthal H-plane logarithmic power ratio slope in dB formed from antenna elements squinted by 90° does not vary in a linear fashion with AOA. In fact the sensitivity of the power ratio varies with AOA from a quasi-linear region from 20° to 70° to a highly non-linear region around either bore sight. The power ratio between antenna elements could be adjusted so that it varies in a more linear fashion with AOA by changing the squint angle between antenna elements. For example by reducing the squint angle from 90° to 72° greater linearity can be achieved in the power ratio over the full range of angle of arrivals. At the same time this creates greater sensitivity to a change in AOA and therefore helps mitigate multipath. It should be noted there will also be cross-over loss associated with a lower squint angle of 72° since the H-plane beamwidth is 78° as seen in Figure 4.23. The cross-over loss implies that the signals in each receiver channel used to form the power ratio are now larger due to a higher gain for a given AOA. Thus the SNR will be larger leading to a lower angular error. Ultimately there is a tradeoff between the number of receive channels and the resolution. As a compromise a prototyped array of four receive antenna elements was chosen.

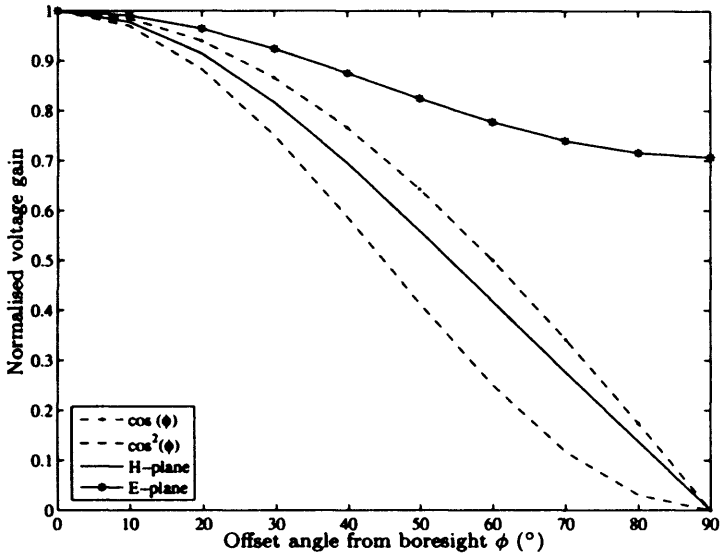


Figure 4.23: Field pattern comparison between the modeled E and H planes and the normalised functions $cos^2(\phi)$, $cos(\phi)$

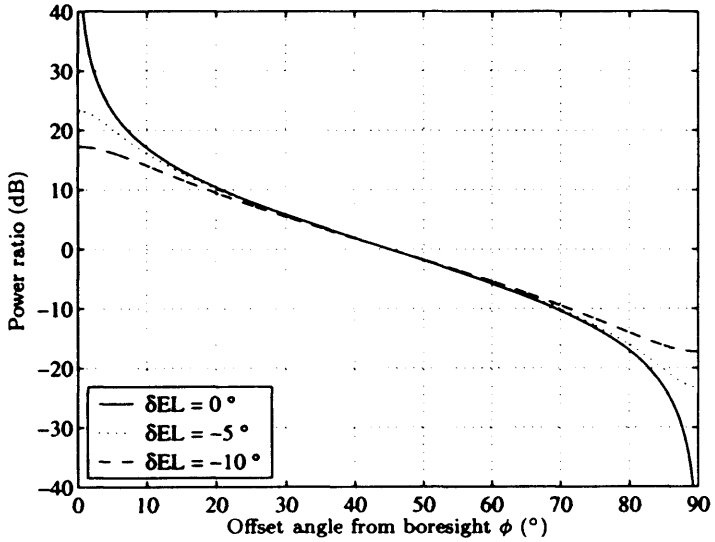


Figure 4.24: Azimuthal only H-plane power ratio compared with the power ratio recalculated for two elevation angular offsets

4.3.2 Array accuracy

The effect of channel imbalance and noise on the direction finding performance of the array arrangement are now investigated. To keep the analysis tractable for the moment the H-plane beam pattern of the rectangular path is approximated by a field pattern, $\cos(\phi)$, with a 90° beamwidth. Assuming that the four antennas are suitably deployed at a squint angle of 90° on the reader housing and their outputs detected and logarithmically amplified by the AGC loop, it is possible to develop the power ratio in dB between the pair of adjacent receivers with the strongest received signal strengths,

$$P_{ratio}\,(\text{dB}) \;=\; 10\log_{10}\frac{S_1}{S_2} \tag{4.14}$$

This can be rearranged to provide a voltage ratio in dB and is given by

$$V_{ratio}\,(\text{dB}) \;=\; 20\log_{10}\frac{V_1}{V_2} \tag{4.15}$$

The voltage in each channel is composed of the field pattern that is measured in $\phi°$ relative to the boresight of the antenna element on channel 1 and the receiver voltage gains, denoted by K_1 and K_2. It is given by

$$V_{ratio}\,(\text{dB}) \;=\; 20\log_{10}\frac{K_1\sin\phi}{\mathrm{K_2}\cos\phi} \tag{4.16}$$

The channel balance is represented by K_1/K_2 and is equal to 1 for balanced channels. Replacing this gain ratio by the factor X, the voltage ratio in dB is now,

$$V_{ratio}\,(\text{dB}) \;=\; 20\log_{10}\left(X\frac{\sin\phi}{\cos\phi}\right) \tag{4.17}$$

which can be furthered reduced to,

$$V_{ratio}\,(\text{dB}) \;=\; 20\log_{10}\left(X\tan\phi\right) \tag{4.18}$$

The use of the cosine pattern approximation gives rise to the need to balance carefully the gain of the two channels, represented by the channel unbalance X in dB between the channels as shown in Figure 4.25. This balance is hard to achieve since

the antenna gain of any one lobe varies by only 3 dB over a 45 degree sector, the articulation being only 6.7×10^{-2} dB per degree. We proceed by analytically determining the effect of channel unbalance on the angular error. The actual error free angle of arrival for balanced channels with X=1 is given by

$$\phi = \tan^{-1}(V_{ratio}) \tag{4.19}$$

Consequently the angular error in the measurement of the angle of arrival for unbalanced channels is therefore

$$\varepsilon = \tan^{-1}(X \tan\phi) - \phi \tag{4.20}$$

To determine the worst case error, the angular error ϵ in Equation 4.20 is differentiated with respect to ϕ and then set to zero,

$$\frac{\delta\varepsilon}{\delta\phi} = \frac{X\sec^2\phi - 1 - X^2\tan^2\phi}{1 + X^2\tan^2\phi} = 0 \tag{4.21}$$

By setting the numerator to zero and solving for the roots we find the solution as

$$X = \frac{1}{\tan^2\phi}$$

rearranging this we find the worst case error at

$$\phi = \tan^{-1}\frac{1}{\sqrt{X}}$$

This is substituted into Equation 4.20 to form the angular error, ε in terms of the channel unbalance, X,

$$\varepsilon = \tan^{-1}\sqrt{X} - \tan^{-1}\frac{1}{\sqrt{X}} \tag{4.22}$$

A plot of decibel unbalance versus angular error for the cosϕ pattern is shown in Figure 4.26. A 1 dB unbalance in gain, for example, will cause a 3.2 degree peak error as shown. This result is compared with the angular error for a pair of antenna

elements with a $\cos^2\phi$ field pattern similarly squinted by 90° in azimuth. The $\cos^2\phi$ field pattern has an angular error that is half that of the $\cos\phi$ field pattern for a given channel unbalance. This means that the channel unbalance is smaller for a field pattern with a lower beamwidth because of the greater articulation of the power ratio. The articulation of the antenna can be further improved by increasing the cross-over loss. Intuitively this cross-over loss is obtained by lowering the squint angle between antenna elements and consequently increasing the number of the elements covering the 360° field of view.

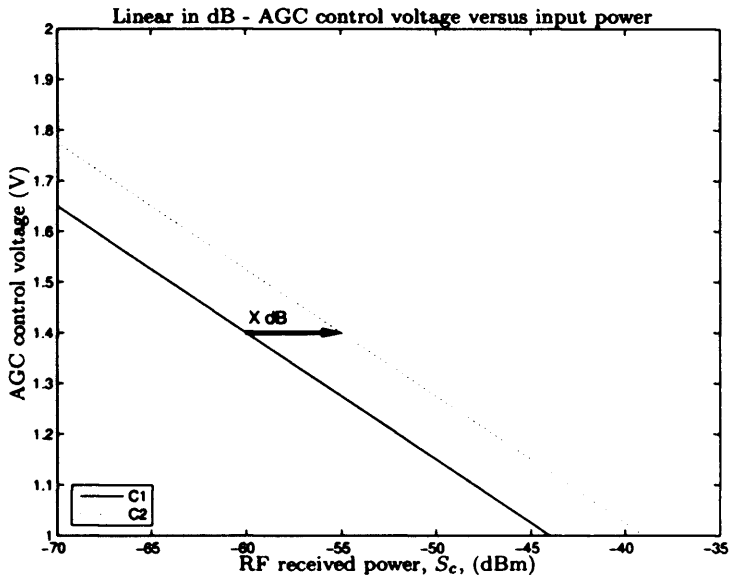


Figure 4.25: Channel unbalance X dB between adjacent receivers channels on the four element reader array

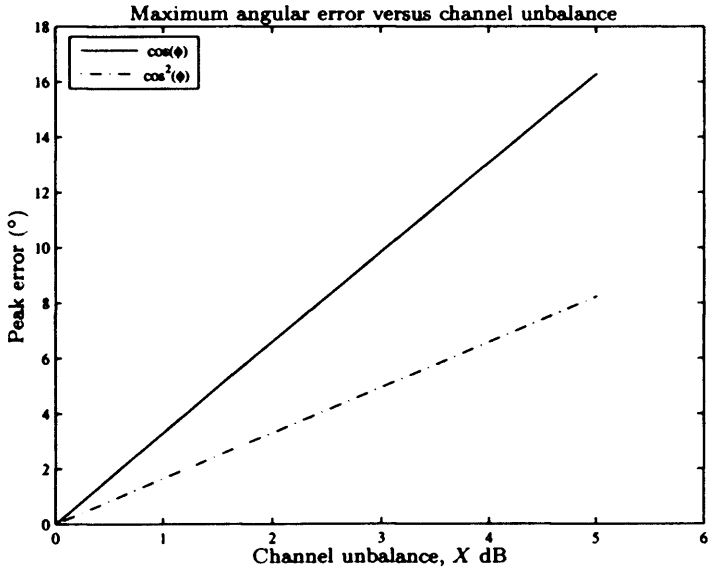


Figure 4.26: The DF angular error versus channel unbalance for the field patterns $\cos^2(\phi)$ and $\cos(\phi)$

A detailed investigation into the effect of squint accuracy and antenna beamwidth on the angular error is not pursued here because the AOA is a non-linear function of $\tan\phi$ and for squint angles $< 90°$, the mathematical differentiation and substitution of compound trigonometric formula is non-trivial. So having considered the effect of channel unbalance on the accuracy of the bearing, we shall proceed by examining the effect of noise on the direction finding performance of the array arrangement [28]. The angular resolution of the direction finding array is affected by the level of noise because the beamwidth of the antenna element is always limited. For instance a system with a desired angular resolution of 1° using a 90° beamwidth antenna element requires a signal to noise ratio of at least

$$10\log_{10}\frac{90}{1} = 19.5 \text{ dB}$$

It is obvious that if no signal integration were applied, 19.5 dB would impose severe sensitivity and dynamic range problems. To analyze the effect of noise on the angular performance of the array we must linearize the power ratio function around the cross-over point. As the voltage ratio in dB restated here for a balanced channel as

$$V_{ratio}\,(\text{dB}) \;=\; 20\log_{10}(\tan\phi)$$

has mixed articulation: high articulation around the boresight (less susceptible to noise errors) and medium articulation around the cross-over (more susceptible to noise errors). We choose to linearise the function by fitting a polynomial to it over the range $\pm 35°$ around the cross-over point. The polynomial is a straight line fit between the voltage ratio in dB and the angular displacement from the cross-over point, Ψ. It is described mathematically as

$$V_{ratio}\,(dB) \;=\; 0.37\Psi \tag{4.23}$$

By the rearrangement of this relationship the angular displacement, Ψ, can be determined from the cross-over point in terms of the estimated channel voltages, V_1

and V_2, of the pair of adjacent receivers channel with strongest signal strength, as

$$\Psi = 54\ (\log_{10}V_1 - \log_{10}V_2) \tag{4.24}$$

The absolute change in the offset angle Ψ with respect to V_1 and V_2 can be obtained by taking the partial derivatives as follows,

$$\begin{aligned} \frac{\delta\Psi}{\delta V_1} &= 7.4\log_{10}(e)\left(\frac{1}{V_1}\right) \\ \frac{\delta\Psi}{\delta V_2} &= 7.4\log_{10}(e)\left(\frac{1}{V_2}\right) \end{aligned} \tag{4.25}$$

The gradients representing the partial derivatives vary with the received voltage as illustrated in Figure 4.27.

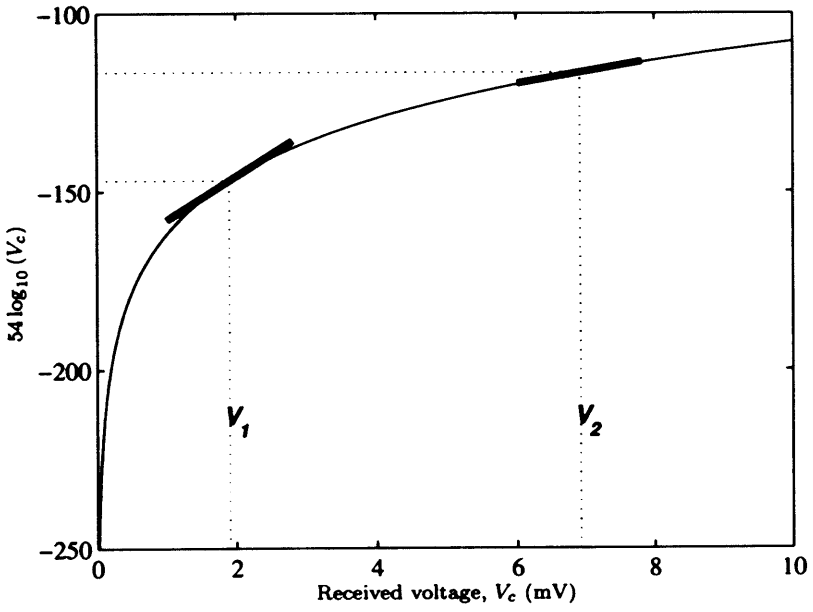


Figure 4.27: Linear plot showing how the slope of the logarithmic function varies with the received voltage, V_c. The curve is plotted for peak voltages corresponding to channel input power levels, S_c, in the range -104 dBm to -30 dBm

The root mean square noise voltage on the input to each channel is, $\frac{\Delta V}{\sqrt{2}}$,

where ΔV is the peak noise voltage. This is transformed to an rms output voltage deviation, $\Delta\Psi$. Solving the Equation 4.25 for Ψ in terms of the rms values,

$$\begin{aligned} \Delta\Psi_1 &= 54\log_{10}(e)\left(\frac{1}{V_1}\right)\frac{\Delta V_1}{\sqrt{2}} \\ \Delta\Psi_2 &= 54\log_{10}(e)\left(\frac{1}{V_2}\right)\frac{\Delta V_2}{\sqrt{2}} \end{aligned} \tag{4.26}$$

The total rms error is found by taking the individual rms errors in both channels and adding them together as the square root of the sum of squares

$$\begin{aligned} \Delta\Psi_{rms} &= \sqrt{\Delta\Psi_1^2+\Delta\Psi_2^2} \\ \Delta\Psi_{rms} &= 23.45\sqrt{\frac{\Delta V_1^2}{2V_1^2}+\frac{\Delta V_2^2}{2V_2^2}} \end{aligned} \tag{4.27}$$

By closer inspection of this equation we note that the noise power represented by

$$\frac{\Delta V_1^2}{2} \quad \text{and} \quad \frac{\Delta V_2^2}{2}$$

are both uncorrelated and can be set to N. In additon V_1^2 and V_2^2 are the channel signal powers. We can therefore reduce Equation 4.27 to

$$\Delta\Psi_{rms} = 23.45\sqrt{\frac{N}{S_1}\left(1+\frac{S_1}{S_2}\right)} \tag{4.28}$$

The power ratio, $\frac{S_1}{S_2}$, can be defined by first re-expressing Equation 4.24 as a power ratio in dB as shown

$$\Psi = 27\log_{10}\left(V_1^2 - V_2^2\right)$$

then rearranging this relationship the power ratio, $\frac{S_1}{S_2}$ can be stated as

$$\frac{S_1}{S_2} = \frac{V_1^2}{V_2^2} = 10^{\frac{\Psi}{27}}$$

After substitution into Equation 4.28, the angular error due to Gaussian thermal

noise can then be found as,

$$\Delta\Psi_{rms} = 23.45\sqrt{\frac{1}{SNR}\left(1+10^{\frac{\Psi}{27}}\right)} \tag{4.29}$$

where SNR is the signal to noise power ratio. The angular error is plotted for 3 SNRs - 10 dB, 20 dB and 30 dB - as displayed in Figure 4.29. These curves permit the determination of the RMS error at each angle of arrival. A general rule of thumb can be used to estimate the angular bearing accuracy. The rms angular error of an antenna element with a given angular beamwidth is inversely proportional to the square root of the SNR. For example, a 20 dB signal to noise ratio is 10:1 and as a consequence an array using a 90 degree beamwidth antenna elements has a 9° average error. From the curves it may be determined that the average DF error for 20 dB signal to noise ratio, of all of the errors, plus and minus, is actually 6°. This is a couple of degrees lower due to the linearisation of the logarithmic slope, V_{ratio}, around the cross-over point. It should also noted that if the received signal has a low SNR, integration of the AGC detection voltage can improve the SNR and thus improve the accuracy of the received signal strength estimates. This is explained further in Appendix B. It should be also be noted that mutual coupling has been discounted from the analysis because of the antenna separations and orientations when mounted on the reader enclosure.

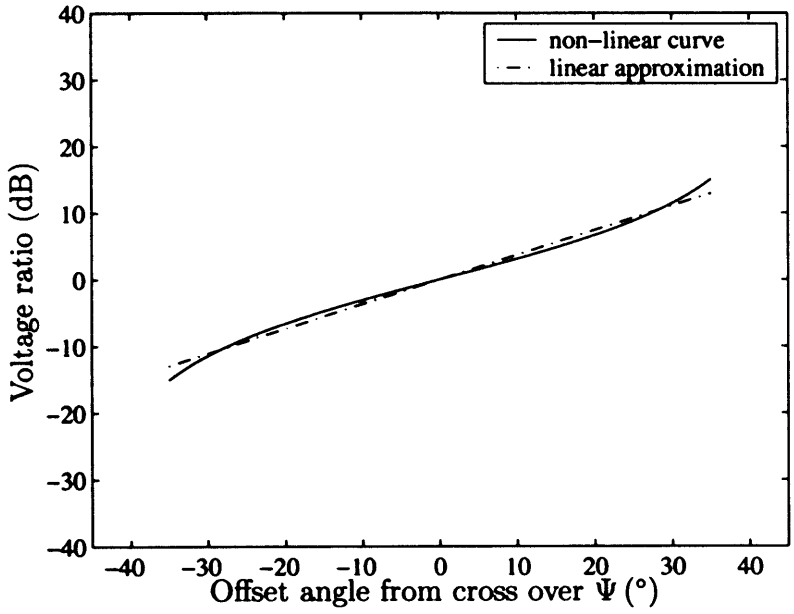


Figure 4.28: Polynomial fit of the logarithmic voltage ratio $\tan(\phi)$ around the cross-over point between the field patterns of adjacent antenna elements of the DF array squinted by 90° in azimuth

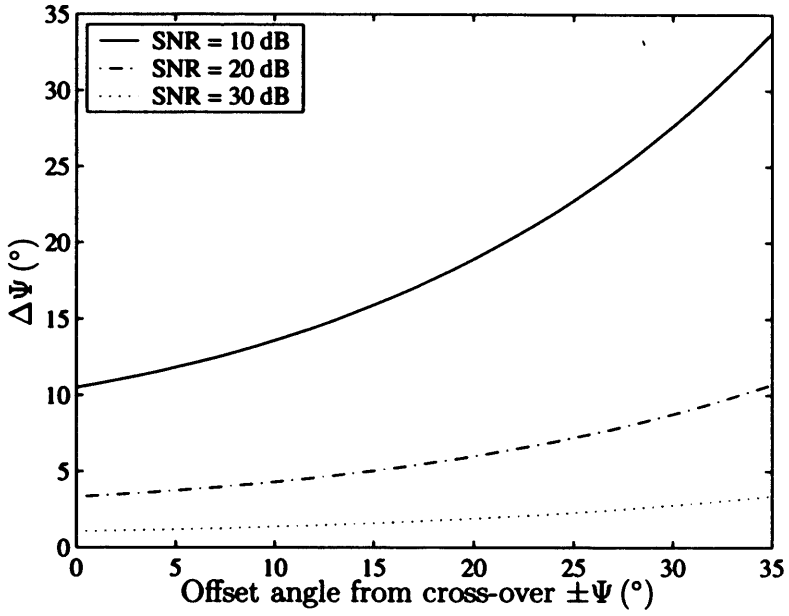


Figure 4.29: The DF angular error versus the angle from the cross-over point, Ψ, plotted for three SNRs

4.3.3 Design and testing of the microstrip patch antenna

The antenna consisted of an inset fed rectangular microstrip patch which is pictured in Figure 4.30. The patch had a simple design because it allowed a quick development cycle since without this component the rest of the transmitter and receiver system could not be completely tested. The performance of the antenna would ideally have a gain greater than 0 dBi, have an S_{11} return loss $\geq$ 25 dB, have a beamwidth of 90° and a roll off consistent with a $10\log_{10}(\cos\phi)$ pattern. This specification presented a significant challenge here because the size of the antenna board and the location of the SMA feed point to the antenna board were restricted by the base station housing. As a result the feed line to the the antenna had a 90° bend to accommodate the H-plane radiation pattern in the azimuth plane and to avoid encroachment on the AGC and RF boards located directly behind the antenna board. This meant that the cross polar components of the radiation pattern were unavoidably much higher than if the patch was fed with a straight feed. This can be seen in the measured polar radiation patterns for the H-plane and E-plane shown in Figures 4.31 and 4.32.

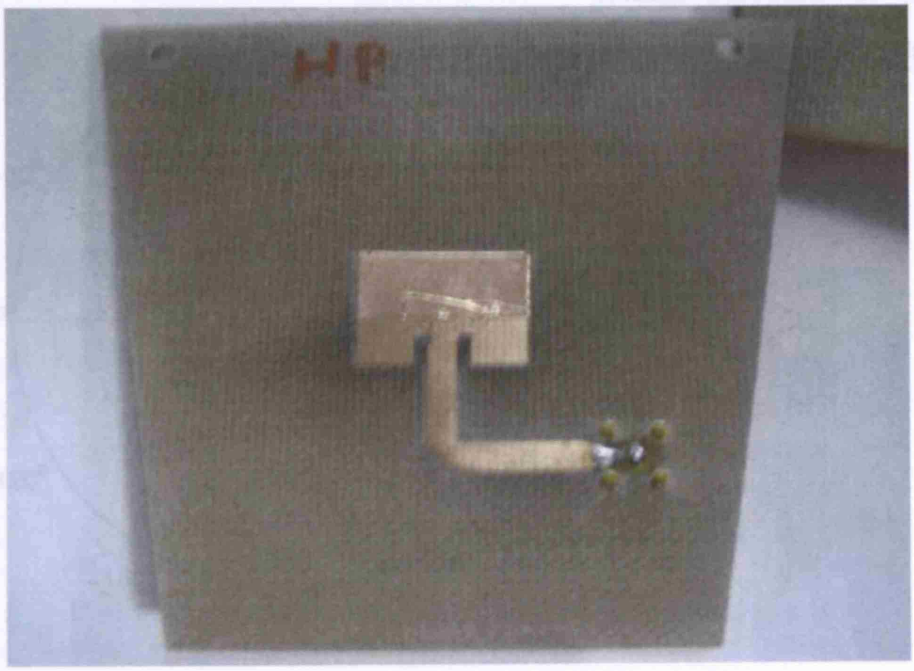


Figure 4.30: Picture of the rectangular inset fed microstrip patch used in the DF array

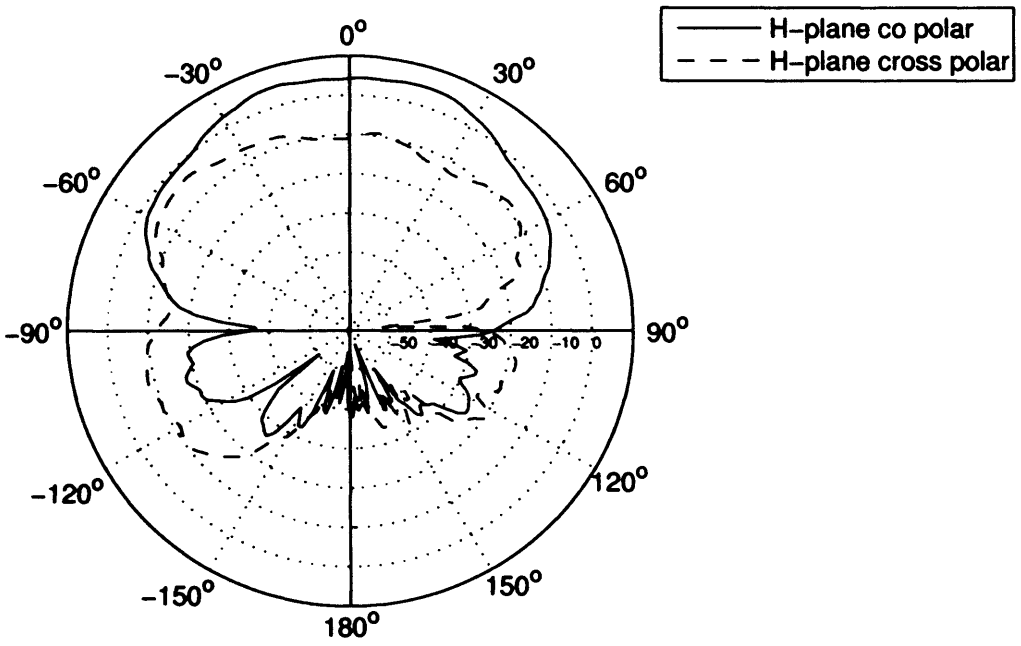


Figure 4.31: Polar H-plane radiation pattern of the patch antenna element measured in the anechoic chamber using a step size of 2°

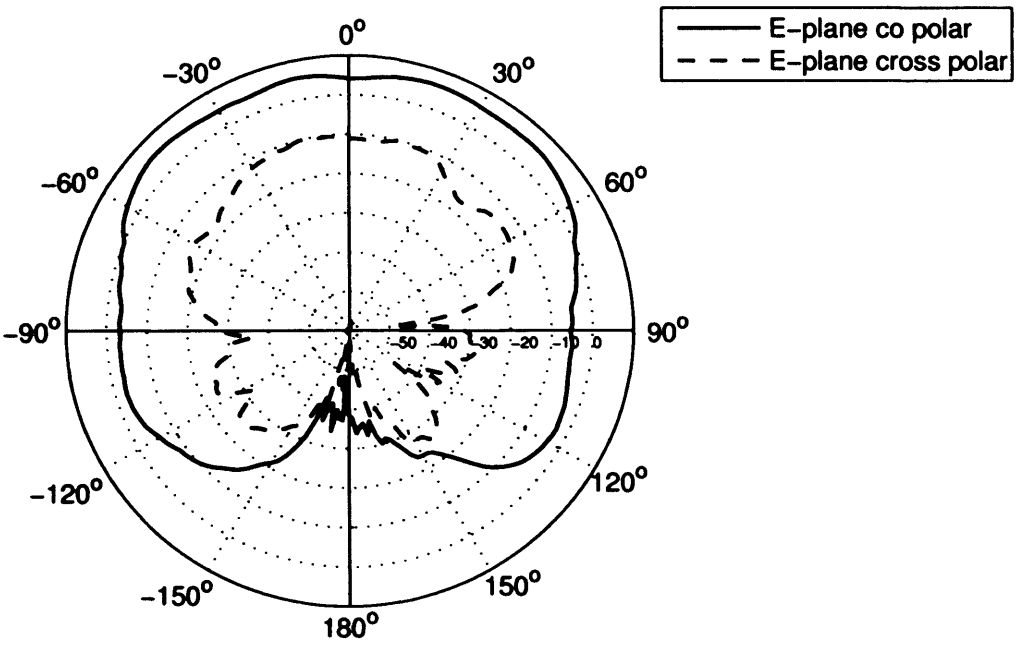


Figure 4.32: Polar E-plane radiation pattern of the patch antenna element measured in the anechoic chamber using a step size of 2°

The cross-polar response adds ambiguity to the measurement of the angle of arrival if the polarization state of the received signal is not within the vertical plane. The measured polar radiation plots were then converted to a linear plot of gain in dB versus angle in degrees as shown in Figure 4.33. The linear plot shows ripple in the measured beam patterns of no more than 1 dB that may be caused by a combination of the step size and the centre of the board may not be perfectly alligned with the centre of rotation. In particular we see that H-plane radiation pattern shows some asymmetry with the feed side radiation a dB or so higher and there is some insensitivity around the boresight. The beamwidth of the H-plane radiation pattern is approximately 80° but varies slightly depending on which power ratio is formed between the asymetrical sides of the beam. This compares to 65° for a $(\cos^2\phi)$ beam pattern and 90° for the $(\cos\phi)$ beam pattern. Even though the measured H-plane radiation pattern does not smoothly rolloff if we plot the logarithmic power ratio in dB as shown in Figure 4.34 we find that the slope is very similar to a $20\log_{10}(\tan\phi)$ function rolloff with offset of less than 1 dB apparent across all angle of arrivals. The $20\log_{10}(\tan^2\phi)$ function is also plotted to show that the beam pattern has not quite met this characteristic. Theoretically it is a better function than the $20\log_{10}(\tan\phi)$ function but would require a number of extra iterations of antenna design, PCB manufacture and testing in the anechoic chamber in order to attain this narrow beamwidth. The $20\log_{10}(\tan\phi)$ function is adequate for the purpose of the experiments.

The completed antenna array on the base station housing is shown in Figure 4.36. Prior to any DF measurements the return loss of each antenna element was also measured. The return loss was measured at 32 dB as illustrated in Figure 4.35 and was achieved by tuning which involved adding sticky copper tape to the length to provide the resonant frequency and then adjusting the inset using a scalpel to obtain the 50 Ohm impedance. Figure 4.35 also shows that the designed antenna element has a wide operating bandwidth of circa 300 MHz at a return loss of 10 dB.

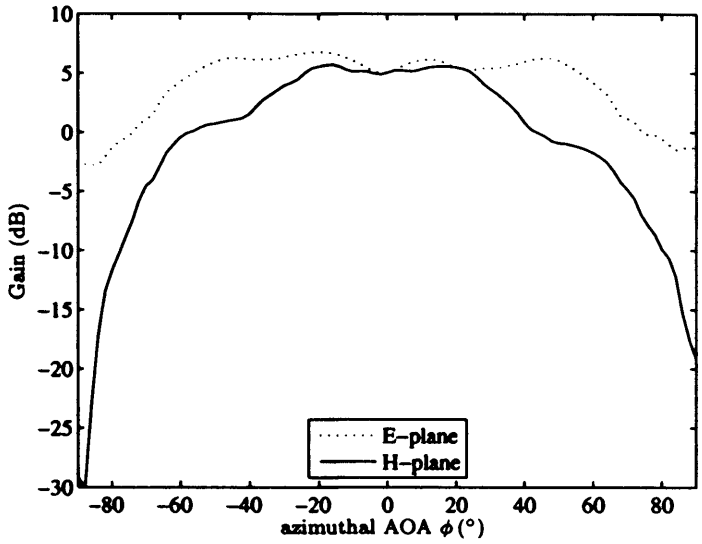


Figure 4.33: Linear radiation plots of elevation E-plane co-polar and azimuthal H-plane co-polar radiation patterns measured in the anechoic chamber using a step size of 2°

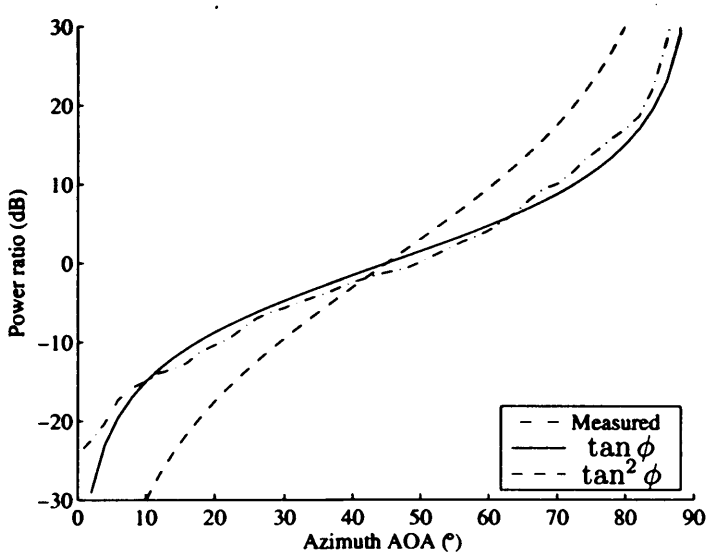


Figure 4.34: Measured aziumuthal voltage ratio between the field patterns of adjacent elements squinted by 90° in aziumth compared with logarithmic voltage ratio functions, $\tan\phi$ and $\tan^2\phi$

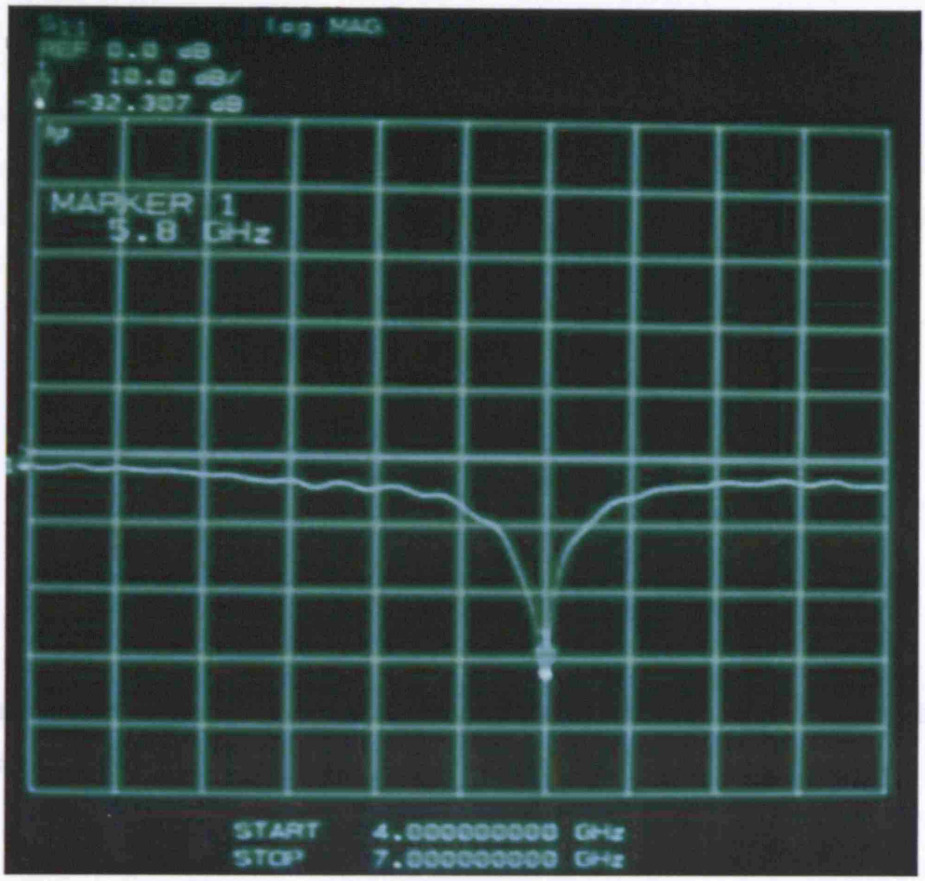


Figure 4.35: Measured antenna element return loss for the completed patch antenna - frequency (300 MHz per division) and return loss (10 dB per division)

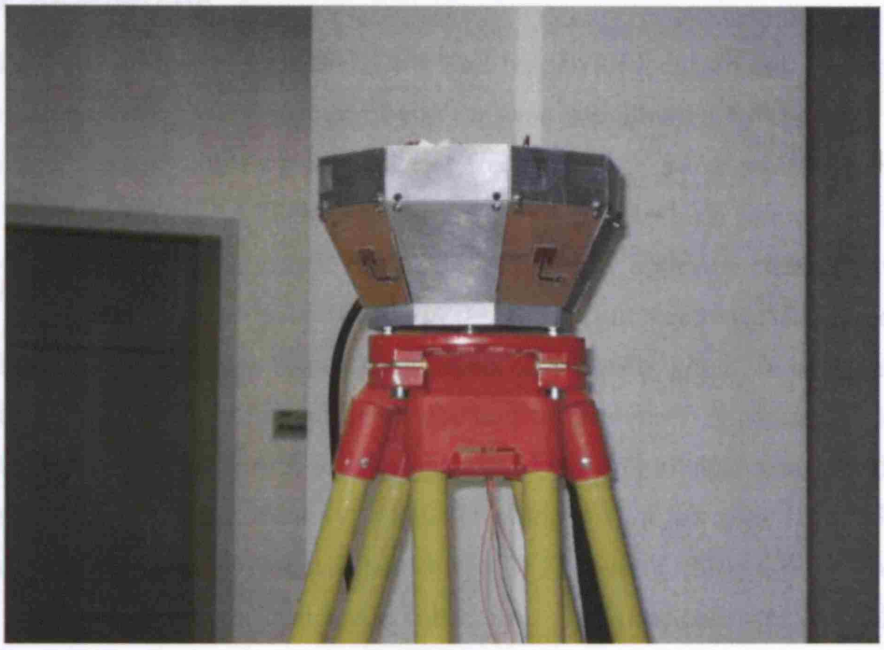


Figure 4.36: The completed four element antenna array used in the reader

4.4 Conclusions

The development of the RF-ID tag and reader for location determination was shown in this chapter. The design of the tag was complicated by the high division ratio caused unavoidably by a high carrier frequency and a low reference frequency used in the synthesiser. A high division ratio can lead to high reference sidebands on the output of the synthesiser. The addition of any spurious component on the output of the synthesiser is of course undesirable, but reference sidebands are a particular nuisance because they fall on adjacent channels which, in communication applications, may give rise to adjacent channel interference. A considerable reduction in the reference sidebands was demonstrated by the judicious use of addditional filtering involving a third order type II filter along with a low pass filter. Since the low pass filter was placed directly before the VCO tuning voltage pin, it also suppressed many other sources of noise in the circuit. The time constant of the low pass filter was swapped with one of the time constants of the third order type II filter as illustrated in Figure 4.2(b) for some additional filtering of the op-amp noise whilst retaining the same loop characteristic. The synthesiser was then carefully modified to allow FSK modulation without disturbing the dynamics of the loop. Thus the output from tag has a clean signal spectrum with a phase noise of -120 $dBcHz^{-1}$ @ 500 kHz offset and there were no spurious signals close to the carrier.

Each reader can also provide an indication of the AOA to the transmitter using the received signal strength indicator (RSSI) that is naturally available from the detection circuitry used in the AGCs on each receiver. Each receiver in the reader has been designed with 80 dB of dynamic range and is linear in dB in the range -30 dBm to -90 dBm. This corresponds to a maximum range of 400 m for identification and location whenever the propagation conditions are favourable. The receiver has a sensitivity of -100 dBm which will potentially allow the identfication only of tranmsitters at ranges in excess of 400 m. The dynamic range of the receiver is controlled by the AGC. The AGC loop was designed with great care otherwise instability may occur in the receiver and the signal strength indicator may not be reliable. In the design chosen here there were no problems with stability as a

compromise was made between fast locking onto the pre-amble of the data burst and stability.

The direction finding properties of an antenna array was also investigated in this chapter. Specifically, this was performed for a passive direction finding array containing four microstrip path antenna elements that were squinted at 90° in azimuth from each other. The regular arrangement allowed the power ratio between adjacent antenna elements to uniquely define the angle of arrival in azimuth. The accuracy of the array was then evaluated for articulation and the angular bearing error with additive Gaussian receiver noise for an antenna element with an $\cos\phi$ field pattern. The angular errors caused by channel unbalance may in large part be removed by calibration or reduced in size by using an antenna element with better articulation, such as a $\cos^2\phi$ field pattern. The angular error varies with SNR, for instance at the cross-over point a SNR of 10 dB produces 10° of error whilst a SNR of 30 dB produces 1°. The SNR can be increased by using a period of integration of the RSSI on each receiver before sampling. Having established the theoretical limits of the direction finding array an inset fed rectangular patch was then designed with a $\cos\phi$ field pattern. The measured co-polar H-plane logaritmic power ratio in dB for the isolated antenna element corresponded well with a $20\log_{10}(\tan\phi)$ function. While the measured power ratio in dB has not been measured between separate antenna elements when mounted on the reader housing it should approximately conform to the expected aziuthmal H-plane response even though the geometry of the housing does contain a base plate that may slightly disturb the far field E-plane radiation pattern.

5. PROPAGATION MODELLING

In this chapter the use of propagation modelling in evaluating the performance of the reader network is demonstrated. This type of modelling provides vital information on how the indoor environment impacts on the triangulation accuracy of the reader(s). For instance, wall placements, corner reflections, furnishings and badly placed readers can cause false bearings of the tag position by a combination of multipath propagation and the lack of a direct line of sight. This is a serious concern. In response to this a ray tracing technique has been developed to model the main reflective features of a typical indoor environment within which the RF signal operates and is applied in progressing degrees of complexity from a two dimensional model to three dimensional model. The three dimensional model is then used to calculate the angular bearing errors and location errors for several readers arranged in a network whenever the transmitter is moved along different paths.

5.1 Ray tracing origin and application

The application of ray tracing in the modeling of rays of light in 3D computer graphics environments is widely known. It is a general technique taken from geometrical optics of modeling the path taken by light by following rays of light as they interact with optical surfaces. It works by tracing, in reverse, a path that could have been taken by a ray of light which would intersect an imaginary camera lens. Similarly, the technique may also be applied to radio propagation at frequencies greater than approximately 900 MHz because radio waves can be described as travelling along localized ray paths which are approximately a straight line. The linear trajectories, which arise from propagating radio waves at high frequency, stems from Maxwells equations [29].

5.1.1 Model introduction

The ray tracing model uses the theory of geometrical optics (GO). This theory is applied whenever the dimensions of the radiating or scattering objects are the order

of many wavelengths. GO only considers direct, reflected and refracted rays. The rays travel in straight lines and the corresponding amplitude, phase and direction following reflections or refractions within a medium can be calculated vectorially using Snell's and Maxwell's equations. All rays contribute to the received radiation field intensity which is found by the sum of all the contributions from each ray path. The impulse response as shown in Equation 5.1 can be calculated for any of the reader receiver channels. This allows a straightforward method of numerically calculating the received signal strength on each channel.

$$h(t) = \sum_{n=1}^{N} A_n \delta(t - \tau_n) \exp(-j\psi) \tag{5.1}$$

The received signal $h(t)$ is formed from N time delayed signal components. Each signal component has an amplitude A_n, a time delay τ_n and a phase shift ψ_n. The phase shift ψ is equal to $\psi_o - \frac{2\pi d}{\lambda}$, where ψ_o is the initial phase shift at the transmitter. Since signal propagation in the indoor environment occurs by line of sight paths, path interactions with reflective surfaces and by diffraction, it is necessary to evaluate propagation theory for these three categories before the presentation of both the 2D and 3D models.

5.1.2 Electric field propagation

The power received from a direct path is related to the distance separating the transmitter and the receiver antenna element. In fact the Friis transmission formula in Equation 5.2 shows the power received P_r, from a ray travelling in free space at a distance r from a transmitter, obeying the inverse square law distance rule.

$$P_r = \frac{P_t G_t G_r \lambda^2}{(4\pi r)^2} \tag{5.2}$$

where P_t = power transmitted ; P_r = power received ; G_r = receiver gain ; G_t = transmitter gain ; and r = separation distance. The equation provides an indication of the maximum range of the transmitter from the receiver under the ideal conditions of line of sight, no multipath components and when it assumed that the transmit and receive antenna polarisation are matched. The polarisation referred to

here is the curve that the tip of the electric field vector, $\boldsymbol{E}$, traces out with time at a fixed point in space and is determined by the antenna geometry and its orientation relative to the observer. If there is a polarisation mismatch then the power versus inverse square distance relationship does not hold. Hence in order to provide an accurate appraisal of signal strength received at each receiver element in the ray tracing model we develop an equation not only taking account of transmit polarisation but also changes in polarisation from any reflections and the polarisation of the receive element. Whenever a ray is transmitted, it has a corresponding polarisation state and power gain related to the direction, $T(\phi, \theta)$ in the local coordinate system. On reception at a receiver antenna element the ray has a corresponding polarisation state and power gain related to the direction $R(\phi, \theta)$ in the local coordinate system. Since each antenna element, whether in transmit or receive mode, has a corresponding polarisation projection $\vec{h}_e$ associated with it then there is a possibility of a polarisation loss between the impinging electric field and the receive antenna element. The received ray polarisation and receive antenna element projection vector can then be used to compute an equation defining polarisation loss factor. Once this is defined the voltage developed on the antenna terminals can then be related to the polarisation loss factor and the antenna factor.

The magnitude of the electric field from a radiating antenna at distance r can be found using the effective isotropic radiated power P_{EIRP} as given by

$$\frac{\delta P}{\delta S} = \frac{P_{EIRP}}{4\pi r^2} = \frac{1}{2\eta} E^2 \Rightarrow E = \frac{1}{r}\sqrt{\frac{\eta P_{EIRP}}{2\pi}} \quad \text{V/m} \quad \text{E}_{\text{rms}} = \frac{E}{\sqrt{2}} \tag{5.3}$$

where η_0 is the characteristic impedance of free space which is equal to 377 Ω. The electric field vector, $\boldsymbol{E_{inc}}$, at a distance r has a magnitude of E and complex phase constant e^{-jBr}. This electric field $\boldsymbol{E_{inc}}$ is then received by an antenna element. This receive antenna element has a polarisation projection vector $\vec{h}_e$ which is directly related to $\boldsymbol{F}_\perp$ the transverse part of the radiation vector of the linear patch antenna described in Chapter 4 and is a function of $\hat{\phi}\cos\theta\sin\phi - \hat{\theta}\cos\phi$ [27]. The radiation vector essentially sets the direction of the electric field for a particular viewing direction on the antenna. This vector, $\vec{h}_e$, when formed in a dot product with

the incident complex electric field vector, $\boldsymbol{E_{inc}}$, indicates the loss factor associated between the two polarisation states shown here as

$$p = \frac{\left|\boldsymbol{E_{inc}} \bullet \vec{h}_e\right|^2}{\left|\boldsymbol{E_{inc}}\right|^2 \left|\vec{h}_e\right|^2} \tag{5.4}$$

Conventionally the loss factor is only used whenever there is roll misalignment between antennas with the same polarization. For instance, the loss factor for linear polarization is a function of one variable which is the axis misalignment. However with the three dimensional ray tracing model the loss factor is a function of four variables in (ϕ, θ) between the received signal and the antenna element.

The theory can be completed by deriving an equation which takes into account the effect of polarisation loss factor and the angle of arrival of the ray on the receive element to voltage developed on the antenna output terminals. The normal definition of the effective area of an antenna and the corresponding power gain, $G = \frac{4\pi A}{\lambda^2}$, depend on the assumptions that the receiving antenna is conjugate-matched to its load and that the polarisation of the incident wave matches that of the antenna. The modified area-gain relationship is given by

$$A(\theta, \phi) = e_{load} e_{pol} \left(\frac{\lambda^2}{4\pi}\right) G(\theta, \phi) \tag{5.5}$$

The area-gain relationship is used in the derivation of the antenna factor (AF) which is defined as the ratio of the incident electromagnetic field to the output voltage from the antenna, when terminated in a 50 ohm load. The antenna factor is given by

$$AF(\theta, \phi) = \frac{9.73}{\lambda \sqrt{G(\theta, \phi)}} \quad \mathrm{m}^{-1} \tag{5.6}$$

As a result the received voltage by the antenna element is found as

$$V_t = \frac{PLF \left|\boldsymbol{E_{inc}}\right|}{AF(\theta, \phi)} \tag{5.7}$$

The received voltage is then affected by Gaussian receiver noise prior to the measurement of the bearing between the strongest adjacent receiver channels on each reader. Additive Gaussian noise has a large impact on the bearing calculation whenever the received signal level has a low SNR. A low SNR would occur whenever the transmitter is located at far distances from the receiver, especially near to reflective surfaces like walls.

5.1.3 Dielectric surfaces

All non-perfect conductors are to some extent dielectric and therefore do not reflect electromagnetic waves perfectly. These surfaces therefore have the asssociated parameters of μ (permeability), ϵ (permittivity) and σ (conductivity) respectively. The loss in reflection can be defined as a reflection coefficient, R, that is equal to the reflected electric field divided by the incident electric field, $\frac{\boldsymbol{E_r}}{\boldsymbol{E_i}}$. The reflection coefficient magnitude is always less than one and varies according to the polarization of the incoming wave, the angle of incidence(or grazing angle), the frequency of the wave and the characteristics of the reflecting surface. The reflection coefficients are expressed as $R_\perp$ and $R_{||}$ for the reflected electric field $\boldsymbol{E_r}$ relative to the plane of incidence. $R_\perp$ may alternately also be stated as horizontally polarized with respect to the reflecting surface and $R_{||}$ may also be stated as vertically polarized with respect to the reflecting surface. The interface components are illustrated in Figure 5.1.

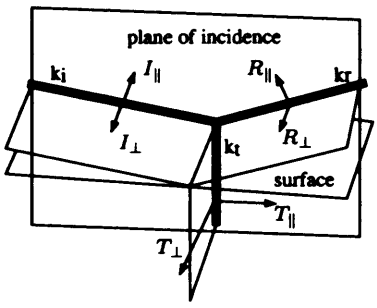


Figure 5.1: An illustration of the reflection interface electric field components.

A couple of assumptions must be stated before deriving the reflection coefficient, R. First, the surface must be approximated by an infinite plane separating the two media so that the surface is the only discontinuity in the environment. Second, both media are homogeneous and the surface between them is perfectly smooth. We begin the deriviation by defining the characteristic or intrinsic wave impedance η as the ratio of electric to magnetic field components, a characteristic of the medium. The intrinsic impedance is defined in terms of μ – the permeability [H/m]and ε – the complex permittivity [F/m] and is given by

$$\eta = \sqrt{\frac{\mu}{\varepsilon - \frac{j\sigma}{\omega}}} \tag{5.8}$$

The imaginary part of the denominator accounts for heat loss in the medium due to damping of the vibrating dipole moments. Both media can be considered as non-magnetic materials in the model and hence $\mu = \mu_0 = 4\pi \times 10^{-7}$ H/m. Provided medium I is where the incident and reflected fields exist and medium II is where the refracted component exists then the square ratio of the two intrinsic impedances used in the derivation can be defined as

$$\eta'^2 = \left(\frac{\eta_1}{\eta_2}\right)^2 = \frac{\varepsilon_2 - \frac{j\sigma_2}{\omega}}{\varepsilon_1 - \frac{j\sigma_1}{\omega}} \tag{5.9}$$

This reduces to the complex permittivity $\varepsilon_2 - \frac{j\sigma_2}{\omega}$ because medium I is assumed to consist of air. The reflection coefficient, which takes account of both amplitude and phase variations in the reflected electric field, can then be formed for $R_\perp$ as Equation 5.10 and $R_\parallel$ as Equation 5.11 [30].

$$\frac{\boldsymbol{E_i}}{\boldsymbol{E_r}} = \frac{\cos\theta_i - \sqrt{n'^2 - \sin^2\theta_i}}{\cos\theta_i + \sqrt{n'^2 - \sin^2\theta_i}} = R_h e^{j\phi_h} \tag{5.10}$$

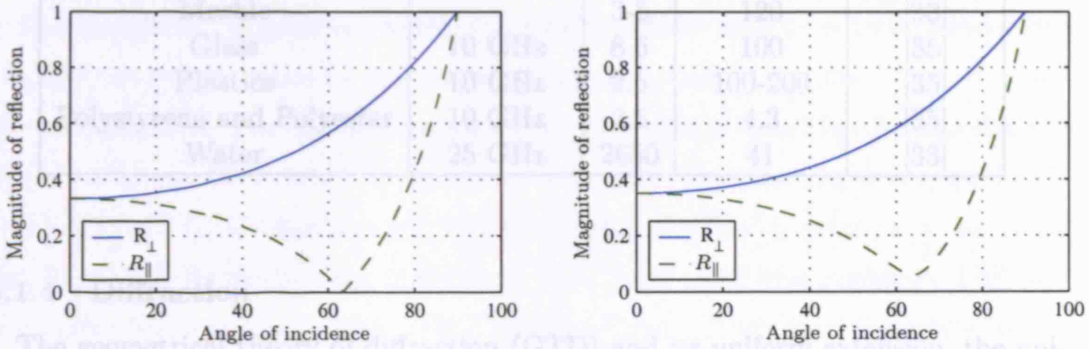
$$\frac{\boldsymbol{E_i}}{\boldsymbol{E_r}} = \frac{n'^2\cos\theta_i - \sqrt{n'^2 - \sin^2\theta_i}}{n'^2\cos\theta_i + \sqrt{n'^2 - \sin^2\theta_i}} = R_v e^{j\phi_v} \tag{5.11}$$

The reflection coefficients for an air-wall reflection are now calculated for two types of material. The results are shown as Figures 5.2(a) and 5.2(b). These graphs show

an accurate description of reflection coefficients using the derived Fresnel coefficients. Interestingly when the angle of incidence is zero, the reflection magnitude for both planes can be approximately given by

$$R = \frac{1 - \sqrt{\varepsilon_r}}{1 + \sqrt{\varepsilon_r}}$$

For dielectric constants, ε_r of 2 and 4 the corresponding to reflection magnitudes are 0.17 and 0.33.



(a) A low loss dielectric with $\varepsilon_r = 4$ and $\sigma = 1e-8$

(b) A lossy dielectric with $\varepsilon_r = 4$ and $\sigma = 0.4$

Figure 5.2: A comparison of the reflection coefficients $R_\perp$ and $R_\parallel$ for two dielectric types

The transmitted coefficients are not considered here because the simplified ray model uses a reflection only model. Furthermore, the analysis in Equations (5.11, 5.10) neglects the effect of the expected internal reflections within the width of a finite structure such as a wall. Nevertheless the reflection coefficient should be sufficiently accurate because the material thickness is large compared to the wavelength[14], for instance a 6 inch block in comparison to the free space wavelength of 5.17 cm at 5.8 GHz. The calculation of the reflection coefficient for reflective surfaces is furthered complicated by the choice of multilayered structures rather than pure dielectrics. For example a wall can consist of many layers each with different characteristics and properties. Consequently wide variations in re-

ported permittivity and conductivity are found for most materials. These values are heavily dependent upon the test frequency but also the composition of the material [31]. Table 5.1 shows some of the measured dielectric properties for some common materials found in the literature.

Table 5.1: Dielectric properties for some common building materials

Material	frequency	ε_r	$\tan\delta \times 10^4$	Reference
Paper/pressboard	1 kHz	3-4	20-100	[32]
Paper	10 GHz	2.6	400	[33]
Wood	1-100 MHz	3	350-1000	[32]
Brick	4 GHz	4-6	unknown	[34]
Marble	-	3-5	120	[33]
Glass	10 GHz	8.6	100	[35]
Plastics	10 GHz	2.5	100-200	[35]
Polystyrene and Polyester	10 GHz	2.5	4.3	[35]
Water	25 GHz	2650	41	[33]

5.1.4 Diffraction

The geometrical theory of diffraction (GTD) and its uniform extension, the uniform GTD (UTD) extend the GO theory by introducing a new type of ray, known as the diffracted ray. Diffraction of an electromagnetic wave is caused by discontinuities in a surface. The discontinuity acts as a radiating point or edge for a small fraction electromagnetic wave allowing for example the phenomenon of diffraction around a corner. It is not used in the reflection model used here because it is felt that at the frequency of 5.8 GHz considered here the effect of diffracted rays may well not be significant, particularly for the ray tracing simulation of an empty 3D room, and thus doesn't warrant further investigation. The application of the geometric optics method, which excludes the diffracted ray paths, provides a simple and effective method for predicting the behaviour of the received rays in the room.

5.2 Model A - 2D model

The angular accuracy of each reader when mounted either in the centre or the corner of a square room is first estimated numerically using a discrete two dimensional ray tracing model. The ray tracing program outputs bearing estimates using the received signal strength on each receiver of the reader. The reader consists of four receiver channels with directive antenna beam patterns arranged at 90° relative to each other. The logaritmic voltage ratio in dB between adjacent antenna elements squinted by 90° is defined as $20\log_{10}(\tan\phi)$ to correspond approximately with the measured co-polar voltage ratio shown in Chapter 4. Since the logarithmic function uniquely defines the azimuth angle, the model is then able to output a measured angle of arrival by choosing the strongest adjacent pair of channels from the four receiver channels. The signal strength on each receiver is found by the complex addition of the received rays in the presence of additive Gaussian noise of noise factor (NF). Typical noise figures range from 0.5 dB for very low noise devices, to 4 to 8 dB. In the RF-ID receiver the noise factor was found to be approximately 4 dB and this noise figure is then added to the model. Another feature of the model is the rays received by the reader are present solely within the room which is enclosed and has a rectangular shape. The number of these rays present in the room is related to the number of calculated ray paths. The ray paths chosen for this model consist of a direct line of sight ray path and four first order image ray paths off the four walls as shown in Figure 5.3.

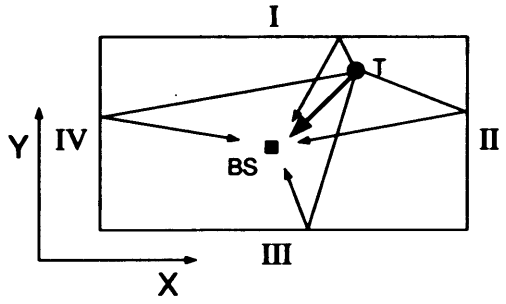


Figure 5.3: An illustration of 2D ray propagation using first order reflections

The model is designed to give the angular bearing errors at any transmitter

position within the room for the following conditions. First, the model assumes that everything is in the same plane, the transmitter, reader and ray paths. There is therefore zero elevation difference between transmitter and reader. Second, the model also assumes that walls are smooth and the amplitude of these specular reflections from the wall are controlled by a wall parameter, R, which allows a defined amount of reflection by the wall. Third, the model does not consider the effect of polarisation loss which has been added to the three dimensional electromagnetic model to be discussed in the next section. Fourth, the room is assumed empty. This means there are no reflective artefacts such as furniture present because the locations of such artefacts are normally not known precisely or unavailable, and if these were considered the complexity of the model is limitless.

5.2.1 Ray tracing image generation

This model requires that the all ray paths are evaluated to determine path lengths, P_L, and receive angles, R(ϕ), when travelling from the transmitter to the receiver elements. The calculation of these parameters involves the use of transmitter and reader ray images off the reflective surfaces. Because the reflecting surfaces used in the model are parallel to the x, y, and z axis, the transmitter images off the walls can be easily calculated because the normal vector to the plane surfaces is the same for all transmitter positions. For instance an image position for a single first order reflection off wall I is calculated simply using the shortest distance between the wall plane and the transmitter position as shown in Figure 5.4. This is significant and alleviates the need in this model to use more sophisticated methods of object images to plane interfaces which are not situated parallel or perpendicular to the co-ordinate system. This theory is shown in Appendix C.2 and would be required for planar reflective surfaces which are defined as a function of more than one axes variable.

To verify that each ray path exists, normally both the base station image and transmitter image are calculated and the paths from these two points are then traced back to the corresponding transmitter and base station positions respectively. Again referring to Figure 5.4, if the intersection point A is the same value for both traces

and it is located within the dimensions of reflection surface then a purely reflected path is said to exist. However in this model the confirmation of point A for both transmitter and base station images is not necessary due to the geometry of the room. If the the geometry of the room was more complicated we would need to check whether the vector from the transmitter image to the base station passes through surface as shown in Appendix C.3. Hence the calculation of one image is then only required and the transmitter image was chosen accordingly for this model.

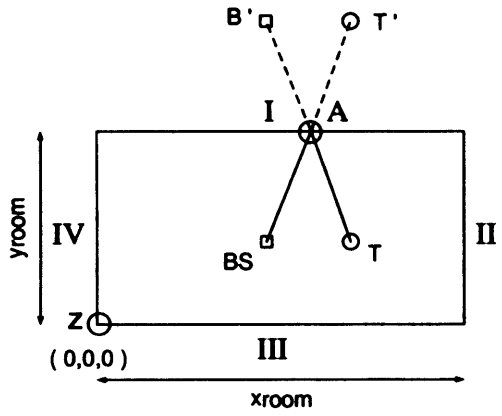


Figure 5.4: Construction of the transmitter and reader reflection images for the first order reflections

The explicit expressions of the transmitter images can be defined in terms of the length and width of the room. If we define the bottom left hand corner as the origin then the image positions for the first order reflections from each wall can be calculated as functions of the transmitter position $T(x_t, y_t)$ and the room dimension x_{room} for walls II, IV and y_{room} for walls I, III.

1. wall I, $x' = x_t, y' = 2y_{room} - y_t$
2. wall II, $x' = 2x_{room} - x_t, y' = y_t$
3. wall III, $x' = x_t, y' = -y_t$
4. wall IV, $x' = -x_t, y' = y_t$

These image points are then used to calculate the ray lengths and receive vectors used in the model.

5.2.2 2D model details

The flowchart described in Figure 5.5 shows the sequence of steps involved in calculating the angular bearing error to a number of transmitter positions on a grid. At each transmitter position four signal components are emitted from a 0 dBi omnidirectionally transmit antenna with 25 mW EIRP transmitter power, they are then either received directly or indirectly via a wall at a reader with four receivers with antenna elements aligned with the faces of the room. The computation of the amplitude and phase of the complex received ray at each receiver requires the calculation of each ray path length, P_L, and voltage gain, $V_g(\phi)$. The ray path lengths and received angles, R(ϕ), are first computed and stored in an array. These received angles, R(ϕ), are then used to calculate the received voltage gain at each element using $\cos\phi$ field pattern antenna element. This field pattern is approximately equal to field pattern measured in the anechoic chamber. Before the voltage gain is computed the received angle of arrival R(ϕ) in the room coordinate system is converted to the local element coordinate system by a simple transformation involving an Euler angle rotation, Ψ_z, about the Z axis. The received complex signal at each receiver is formed by the complex superpositon of the signal component. Assuming a 50 ohm receive system each received ray can be converted into an incident voltage using the antenna factor described in Section 5.1.2 and is proportional to the inverse of the voltage gain, $V_g(\phi)$. The incident voltage is shown in the general form in Equation 5.12 for a direct component and Equation 5.13 for an indirect components containing a magnitude only correction, R, for the reflection off the surface.

$$V_r = \frac{E_r}{AF(\phi)} \exp\left(\frac{iwP_L}{c}\right) \tag{5.12}$$

$$V_r = R\frac{E_r}{AF(\phi)} \exp\left(\frac{iwP_L}{c}\right) \tag{5.13}$$

where V_r = complex receiver voltage, $AF(\phi)$ = Antenna factor, E_r = magnitude of dielectric field for a path length, P_L, w = frequency phasor, $2\pi f$, R = wall reflection constant for indirect rays only, and c = speed of light. The complex received voltages then have bandpass gaussian noise samples added before the modulus is taken so that signal strength estimates can be processed by the adjacent strongest receiver

algorithm. The signal strengths from the two strongest receivers are then used to calculate a measured bearing using a lookup table containing the logarithmic voltage ratio $20\log_{10}(\tan\phi)$ versus azimuthal angle of arrival. The measured bearing is then compared against the actual bearing by first converting both bearings into complex form and then multiplying the actual bearing by the complex conjugate of the measured bearing. The net result is the angular bearing error measured by the reader. This is recorded and the process continues by examining the next transmitter position until the transmitter grid is complete.

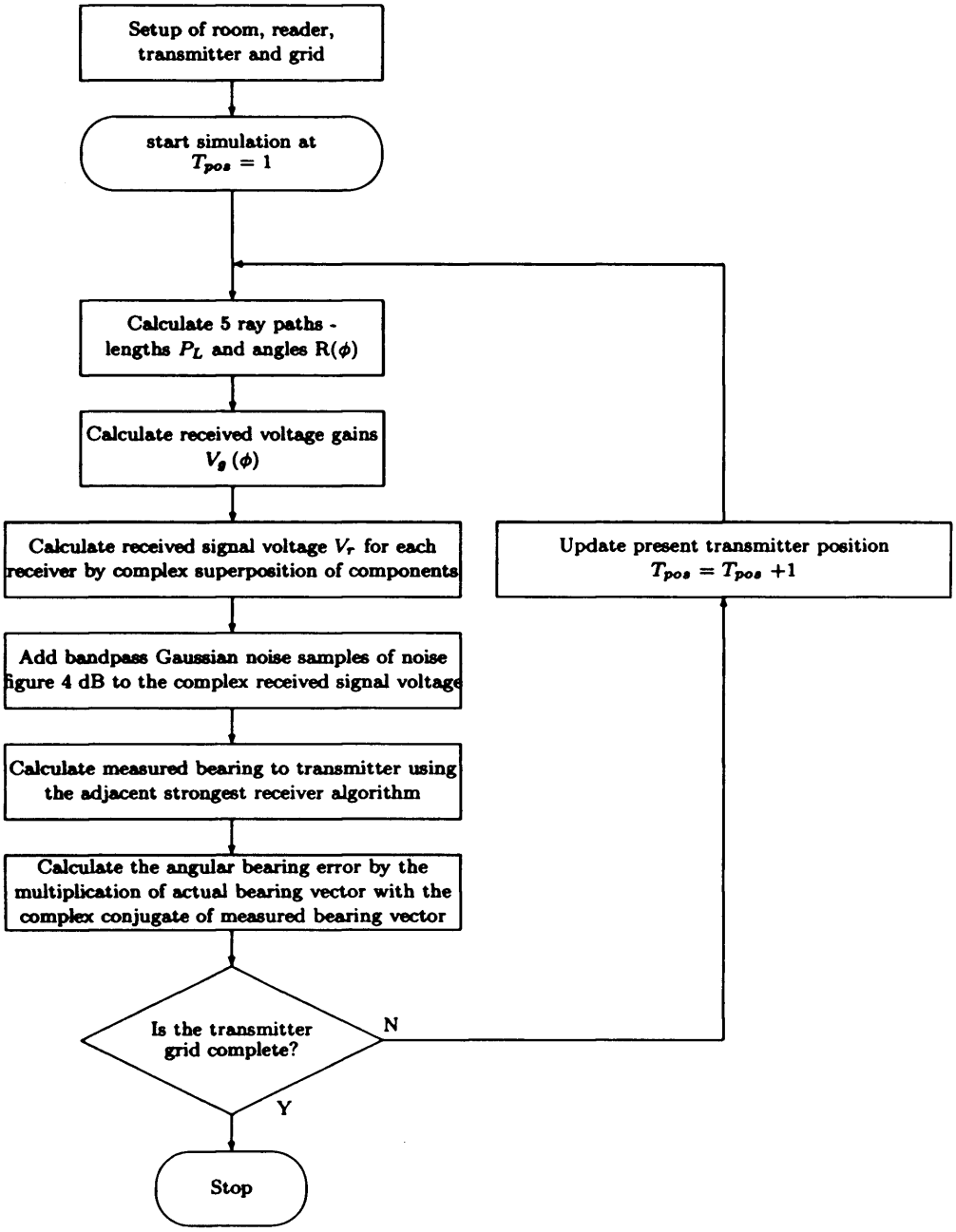


Figure 5.5: Program sequence for model A

5.2.3 Results and discussion on the angular bearing accuracy of a single reader using model A

The simulation can be performed in two ways. The receiver or transmitter can be moved along a given path or over the whole (or part of) the enclosed area. For the results to be presented here, the transmitter was moved in discrete steps of 2.5 m on a grid in x and y around a 25 m square room for a total of 625 transmitter positions. At each transmitter position, the angular bearing error was recorded for a fixed wall reflection magnitude, R, for a stationary reader placed either in the corner or in the centre of the room. The room size and wall reflection magnitude, R, were varied and the measurements of angular bearing for the grid of transmitter positions was repeated. Specifically, the room sizes examined by the simulation were 25 m square and a 12.5 m square with the grid step size scaled for the smaller room sizes to maintain 625 transmitter positions. Similarly, the wall reflection constant, R, was examined for three different values: 1 (the worst case of 100 % wall reflectivity which would occur off metallic surfaces), 0.5 and 0.25 (which corresponds approximately to a dielectric constant, ε_r, in the range 2 - 4 provided we approximate the reflection magnitude at an angle of incidence of 0°. This range of dielectric constants would be expected from materials such as brick, wood and plastics as shown in Table 5.1). The results in Tables (5.2, 5.3) show a comparison of the mean and rms angular bearing errors in degrees for a centre mounted reader and a reader mounted in the bottom left hand corner of the room whenever the model includes a direct line of sight signal (LOS) and also when the model assumes no direct line of sight signal (NLOS).

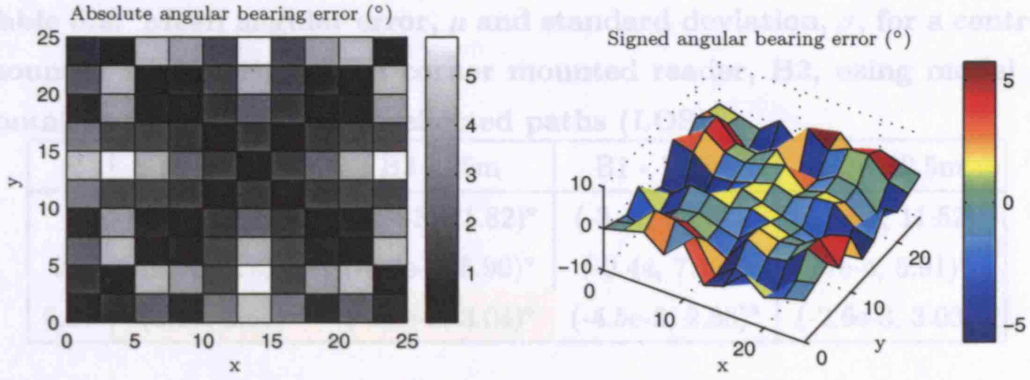
Beginning with the line of sight (LOS) simulation results it can be seen that the lowest errors for both readers mountings occurs along the 45° diagonal between the adjacent antenna elements as seen in Figure 5.7(b) because the reflections from the opposing corner reinforces the directly received signal component. Figures 5.6 and 5.7 each show two alternate illustration of the angular bearing errors for the direct line of sight scenario whenever $R = 25$ % reflectivity. The tabulated results in Table 5.2 from each mounting also show a mean very near zero degrees in nearly

all instances which suggests no bias has been introduced to the calculation of the actual angular bearing from the first order multipath reflected signal components. Furthermore, the results show that both mountings providing similar standard deviation at low levels of reflectivity, though the corner mounting tends to provide better accuracy at reflectivity levels above $R = 0.5$. The rms angular ranges from around 3 degrees at 25 % reflectivity to 6 degrees at 50 % reflectivity. This equates approximately to a location error which varies from 0.52 m to 1.05 m error at an operating range of 10 m.

The independence of reader mounting position must be evaluated further by repeating the simulation only this time without a line of sight received ray. The NLOS results are shown Table 5.3 and the colour map of the spot angular bearing errors shown in Figures 5.8 and 5.9. The corner mounted reader is more accurate than the centre mounted reader with a mean angular bearing error around 0.37° and standard deviation of 18° which does not change significantly whenever the amplitude of the reflections, R, is varied, as is expected since the angular error of the corner mounted reader can be no more than 90°. This means the lack of a LOS signal does not necessarily mean that the angular bearing errors received by the corner mounted reader will be uniformly distributed in the range ±90° because the standard deviation, σ, of 18° is much lower than the expected standard deviation of $90/sqrt(3) = 51.96°$ from angular errors that are uniformly distributed in the range ± 90°. The maximum angular error for the corner mounted reader is related to the number of rays and their interaction rather than their amplitude as seen in column three of Table 5.3. There the magnitude of the reflection coefficient, R, does not change the resulting mean or standard deviation of the angular bearing errors because each multipath component suffers an equivalent reflection magnitude. The angular bearing errors illustrated in Figure 5.9(a) for the corner mounting are smaller whenever the transmitter is positioned on the diagonal between the adjacent pair of corner antenna channels and at the furthest distance away from the walls. Although, there are large angular errors on or near the walls adjacent to the reader. On average the location error from this corner mounted reader is approximately 1.62 m error

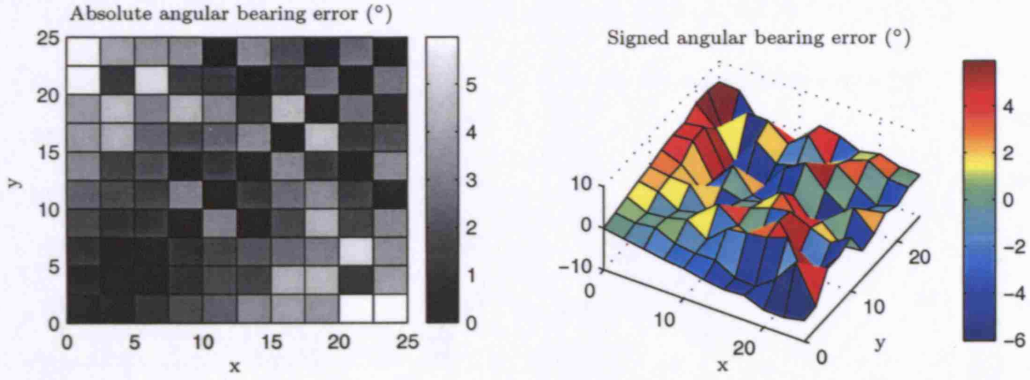
at 5 m range. The centre mounted reader also has a symmetrically error profile as displayed in Figure 5.8(a) where the maximum error occurs on the boresight of each antenna element. The rms angular error tends to be larger than that of the corner mounted reader which implies that reader in this configuration suffers alot from quadrant errors. Quadrant ambiguities can cause angular bearing errors of more than $\pm 180°$ with the central reader mounting and this is twice that of the largest possible angular errors from the corner reader mounting.

These results are very encouraging for accurate indoor location and have indicated the corner reader mounting provides the lowest angular bearing error. Sub-metre location error for ranges greater than 10 m are expected when a line of sight exists whilst a 1.6 metre location error for a range of 5 m are expected whenever there is no line of sight. The results also suggest that bearing error is dependent on reader mounting and that the readers must be optimally placed for the lowest angular error. However the results need to intrepreted with caution since this is a simple model. In order validate these results an upgraded 3D electromagnetic model is required that takes account of the changes, in polarisation on reflection and on reception, on the angular bearing error.



(a) plan view (b) view at 30° azimuth and 60° elevation

Figure 5.6: Plots of the angular error for the central mounted reader B1(12.5,12.5) using model A with both direct and reflected paths (LOS)

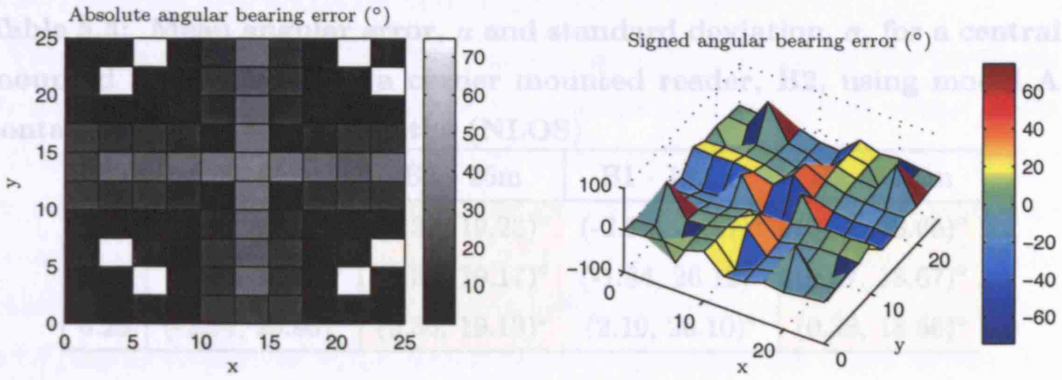


(a) plan view (b) view at 30° azimuth and 60° elevation

Figure 5.7: Plots of the angular error for the corner mounted reader B2(0,0) using model A with both direct and reflected paths (LOS)

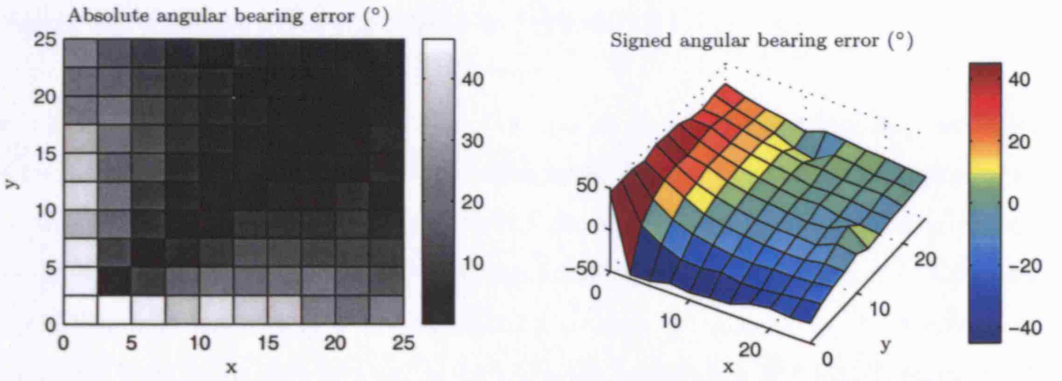
Table 5.2: Mean angular error, μ and standard deviation, σ, for a central mounted reader, B1, and a corner mounted reader, B2, using model A containing both direct and reflected paths (LOS)

R	B1 - 25m	B2 - 25m	B1 - 12.5m	B2 - 12.5m
1	(-0.31, 16.61)°	(3.2e-3, 11.82)°	(-3.41, 34.65)°	(1.4e-3, 11.52)°
0.5	(0.09, 5.73)°	(-4.5e-3, 5.90)°	(-0.44, 7.17)°	(7e-4, 5.91)°
0.25	(0.16, 2.81)°	(-4.8e-3, 3.04)°	(-4.5e-2, 2.86)°	(-2.6e-3, 3.03)°



(a) plan view (b) view at 30° azimuth and 60° elevation

Figure 5.8: Plots of the angular error for the central mounted reader B1(12.5,12.5) using model A with reflection only paths (NLOS) and $R = 25$ % reflectivity



(a) plan view (b) view at 30° azimuth and 60° elevation

Figure 5.9: Plots of the angular error for the corner mounted reader B2(0,0) using model A with reflection only paths (NLOS) and $R = 25$ % reflectivity

Table 5.3: Mean angular error, μ and standard deviation, σ, for a central mounted reader, B1, and a corner mounted reader, B2, using model A containing reflection only paths (NLOS)

R	B1 - 25m	B2 - 25m	B1 - 12.5m	B2 - 12.5m
1	(-0.32, 26.07)°	(0.37, 19.23)°	(-2.86, 28.47)°	(0.37, 18.66)°
0.5	(-2.37, 28.42)°	(0.39, 19.17)°	(-1.84, 26.12)°	(0.37, 18.67)°
0.25	(-2.34, 25.95)°	(0.36, 19.13)°	(2.19, 26.10)°	(0.38, 18.66)°

5.3 Model B - 3D electromagnetic model

This model extends the 2D model by adding greater complexity so that the model accurately describes the actual measurement scenario. The primary change to the previous model is the introduction of ray propagation in 3D because this removes the restriction of the tranmitter and receiver heights being in the same plane. As a consequence, if the reader is mounted at a differing height to the transmitter the model now allows the receiver elements which also have directional beam patterns to be pointed downwards at a chosen range, perhaps at half the expected range for optimum coverage. Since the accuracy of the model is dependent on the number of reflected rays, the received signal on each element is calculated using an additional four 2nd order reflections off opposite walls, as well as the four 1st order reflections defined previously. The model neglects ray paths from floors, ceilings and second order reflections of adjacent walls. This keeps the model simple and gives a reasonable indication of expected angular bearing errors for this scenario. The number of possible ray paths in 3D is considered in Section 5.3.1.

Another equally important update to the model is the application of polarisation loss to each received ray. This loss is computed by the normalised dot product of the polarisation of the received ray with the projection vector of the receive element. The polarisation of the received ray is dependent on the polarisation of the transmitted ray and any changes in polarisation caused by reflections. In addition the model has been improved further by relating the amplitude of the reflections to the relative dielectric constant of the chosen material in the surface rather than a wall reflection magnitude, R. This relative dielectric constant is then used to calculate the reflection coefficients for both the perpendicular and parallel components of the incident electric fields. This represents a more realistic description of the level of reflection off the surface for a particular incidence angle whenever the material of the surface is constructed from a single material. Examples of materials include but are not restricted to metals, where specular reflections can only occur and lossy dielectrics. For instance the sensitivity of the ray tracing model to the wall characteristics could be examined for a low loss dielectric and a lossy dielectric with the

parameters ($\varepsilon_r = 4, \sigma = 1e-8$) and ($\varepsilon_r = 4, \sigma = 0.4$) respectively.

The assumption of an omnidirectional antenna field pattern used in the 2D model does not represent the best description of the antenna element used on the transmitter and is replaced by a more realistic directional field pattern. The tag developed for Optag is worn around the neck using a lanyard and contains the microstrip patch antenna design which was designed for the receiver. Consequently the model contains an antenna element with an $\cos\phi$ field pattern, a 90° beamwidth and is polarised in the vertical direction. The antenna element can be positioned at any fixed height, but is chosen in the model as 1.2 m, and can also point in any direction in azimuth. At the same time, this model retains many of the features of the previous 2D model. The propagating rays are considered only for an enclosed area that is empty and rectangular. There is no consideration of people attenuating the signal or internal objects causing reflections. The model examines the effect of only a *finite* number of ray paths specifically direct rays, single reflected rays off walls, and double reflected rays off opposite walls.

5.3.1 Ray propagation in a cuboid

The number of possible propagation paths from the transmitter to the receiver can be generated using an image tree. This process is implemented by considering all the various surface reflections and transmissions that are possible in the enclosed area. The enclosed area considered here is an empty cuboid which contains the following the number of purely reflected paths p_n, of n reflections and is given by

$$p_n = \begin{cases} 2n^2 + 2 & \text{for } n > 0 \\ 1 & \text{for } n = 0 \end{cases} \tag{5.14}$$

Since this room has a simple shape we are then able to calculate the cumulative sum of the possible paths up to and including n reflection paths can be described as

$$p_{cs} = \frac{4}{3}n^3 + 2n^2 + \frac{8}{3}n + 1 \tag{5.15}$$

The number of true 3D paths, p_n, at order n and the cumulative sum, p_{cs}, of the true 3D paths upto order n, are tabulated in Table 5.4. Also shown for comparison is the number of 3D permutations, p_{perm}, that takes no consideration of non possible reflection combinations and uses the number of surfaces which in this case is six – representing four walls, a ceiling and a floor.

Table 5.4: Ray path permutations for a cuboid

n	p_{perm}	p_n	p_{cs}
0=LOS	1	1	1
1	6	6	7
2	36	18	25
3	216	38	63
4	1296	66	129
5	7776	102	231
6	46656	146	377
7	279936	198	575

Since image generation is an infinite process, a maximum reflection order must be chosen to limit the number of propagation paths considered. A maximum reflection order of 2 has been chosen again for this model because third order rays should be sufficiently attenuated upon reception and removes the need for the calculation of additional ray paths which have negligible effect on the signal received at the antenna elements. Referring to Table 5.4, a 3D model with maximum reflection order of 2 contains 38 possible reflection paths and 25 actual pure reflected paths. Of these 25 pure reflected paths the current model examines the effect of only a finite number of reflected ray paths specifically single reflected rays off walls, and double reflected rays off opposite walls. This represents a total of 8 reflections with the remaining 17 reflections arising from double reflections off adjacent walls and reflections involving the ceiling and floor not included in the model.

The calculation of the path length, transmit and receiver vectors of the first order and second order reflections is performed using the image technique discussed in Section 5.2.1. The explicit expressions of the transmitter images, shown in Figure 5.10, can be defined in terms of the length and width of the room. If we define the

bottom left hand corner as the origin then the image positions for the first order reflections from each wall can be calculated as functions of the transmitter position $T(x_t, y_t, z_t)$ and the room dimensions x_{room}, y_{room}.

1. wall I, $x' = x_t, y' = 2y_{room} - y_t, z' = z_t$
2. wall II, $x' = 2x_{room} - x_t, y' = y_t, z' = z_t$
3. wall III, $x' = x_t, y' = -y_t, z' = z_t$
4. wall IV, $x' = -x_t, y' = y_t, z' = z_t$
5. wall combination (I,III), $x'' = x_t, y'' = y_t - 2y_{room}, z'' = z_t$
6. wall combination (III,I), $x'' = x_t, y'' = 2y_{room} + y_t, z'' = z_t$
7. wall combination (II,IV), $x'' = x_t - 2x_{room}, y'' = y_t, z'' = z_t$
8. wall combination (IV,II), $x'' = 2x_{room} + x_t, y'' = y_t, z'' = z_t$

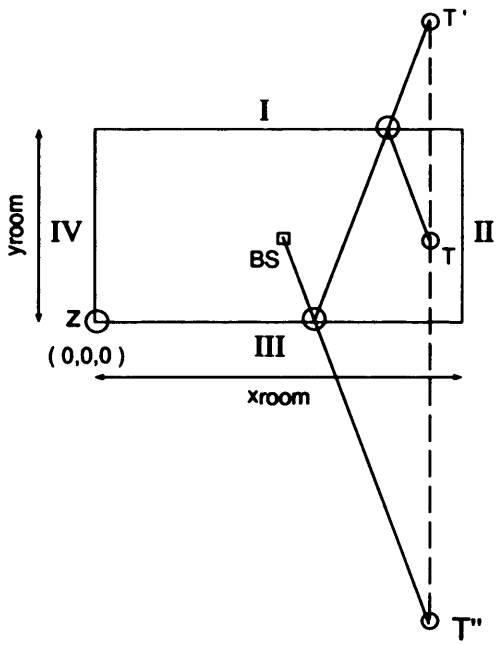


Figure 5.10: Construction of the 1st and 2nd order transmitter images

5.3.2 3D model details

The flowchart described in Figure 5.12 shows the sequence of steps involved in calculating the angular bearing error for a number of transmitter positions on a grid using the 3D model. Part of the initial setup is the definition of the transmitter height and pointing angle in azimuth as well as the reader location and height and receiver element tilt angle in elevation. The simulation then proceeds at each transmitter position to emit nine rays from a directional transmit antenna with 25 mW EIRP maximum transmitter power. These rays are then either received directly or indirectly via a reflecting surface to a reader with four receivers with antenna elements aligned in aziumth with the faces of the room. The path length, P_L, the transmit angle, T(ϕ, θ), and the receive angles, R(ϕ, θ), of each ray is computed using the transmitter and base images and then stored in an array. These angles are then used to calculate the transmit and received voltage gain at each element using the square root of the normalized $\cos(\phi, \theta)$ field pattern of the directional antenna element. Before the voltage gain is computed the angle which is in the room coordinate system is converted to the local element coordinate system by a simple transformation involving an Euler angle rotation, Ψ_z, about the Z axis. This is described in further detail in Appendix C.1. The images generated earlier are used to calculate the transmit and receive vectors that are used later in the model in the calculation of changes in the electric field upon surface reflection. For each surface reflection, the reflected electric field is calculated using the angle of incidence and the polarisation of the incident field with respect to the surface. The reflections take account both of the phase and amplitude of the electric field. The reflections can described analytically in the follow manner. First, the angle of incidence, θ_i, can be defined as

$$\theta_i = \cos^{-1}\left(\frac{\boldsymbol{u_i} \bullet \boldsymbol{n}}{|\boldsymbol{u_i}|}\right) \tag{5.16}$$

where n = surface normal, u_i = vector in the direction of propagation and $\bullet$ denotes the vector dot product. The direction of the horizontally polarized component of the incident electric field can be given as

$$\boldsymbol{h} = \frac{\boldsymbol{u_i} \times \boldsymbol{n}}{|\boldsymbol{u_i}|} \tag{5.17}$$

where $\times$ = the vector cross product. The direction of the vertically polarized component can also be given as

$$\boldsymbol{v_i} = \frac{\boldsymbol{h} \times \boldsymbol{u_i}}{|\boldsymbol{u_i}|} \tag{5.18}$$

These directions components for a ray reflection off a surface whenever the ray travels on a plane perpendicular to that reflecting surface are illustrated in Figure 5.11.

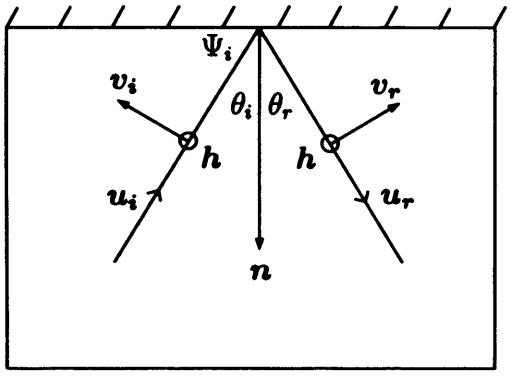


Figure 5.11: An illustration of the change in the polarisation of the electric field before and after reflection

The perpendicular components, $\boldsymbol{E_h}$, $\boldsymbol{E_v}$, of the incident electric field, $\boldsymbol{E_i}$ are then computed as follows.

$$\begin{aligned} \boldsymbol{E_h} &= \boldsymbol{h} \bullet \boldsymbol{E_i} \\ \boldsymbol{E_v} &= \boldsymbol{v_i} \bullet \boldsymbol{E_i} \end{aligned} \tag{5.19}$$

The reflected electric field can then be expressed as Equation 5.20. It should be noted that $\boldsymbol{v_i}$ has changed direction on reflection. R_h and R_v are the reflection coefficients given in Equations (5.10, 5.11).

$$\boldsymbol{E_r} = \boldsymbol{E_h} R_h \boldsymbol{h} + \boldsymbol{E_v} R_v \boldsymbol{v_r} \tag{5.20}$$

where $\boldsymbol{v_r}$ is the direction of the vertically polarized component of the reflected field which can be expressed as

$$\boldsymbol{v_r} = (\boldsymbol{n} \bullet \boldsymbol{v_i})\,\boldsymbol{n} - ((\boldsymbol{n} \times \boldsymbol{h}) \bullet \boldsymbol{v_i})\,(\boldsymbol{n} \times \boldsymbol{h}) \tag{5.21}$$

The analysis for a ray containing double reflections can also be expressed in terms of these reflections coefficients. The complex received rays can then be summarised as follows

1. direct ray: $\boldsymbol{E_r}$

2. first order ray: $\boldsymbol{E_r} = \boldsymbol{E_h} R_{h1} \boldsymbol{h} + \boldsymbol{E_v} R_{v1} \boldsymbol{v_r}$

3. second order ray: $\boldsymbol{E_r} = \boldsymbol{E_h} R_{h1} R_{h2} \boldsymbol{h} + \boldsymbol{E_v} R_{v1} R_{v2} \boldsymbol{v_i}$.

Subsequently the direct LOS signal has a received electric field, $\boldsymbol{E_r}$, with a magnitude of $|\boldsymbol{E_r}|$ at a path length, P_L, and can be described by

$$|\boldsymbol{E_r}| = \frac{1}{P_L} \sqrt{\frac{\eta P_{EIRP} G_t(\phi, \theta)}{2\pi}} \tag{5.22}$$

where η_0 = characteristic impedance of free space (377 Ω) and $G_t(\phi, \theta)$ is the normalised power gain of the transmit antenna. The first order and second order rays also have electric fields that vary with the ray path length but are also attenuated further by the reflections (R_{h1}) and (R_{h1}, R_{h2}) respectively.

The next step in the computation is the conversion of the received electric field of each ray to a voltage which requires the calculation of the PLF and AF before the superposition of the rays at each antenna element. The polarisation loss factor is then calculated by the normalised dot product of the incident Electric field vector and the projection vector, which is a function of $\hat{\phi} \cos\theta \sin\phi - \hat{\theta} \cos\phi$, of the receive element at the receiver angle R(ϕ, θ). Each received ray can be converted into an incident voltage using the antenna factor was described in Section 5.1.2 and is proportional to the inverse of the voltage gain, $V_g(\phi)$. The sum of the received voltages is shown in the general form in Equation 5.23.

$$V_r = \frac{PLF(\phi, \theta) |\boldsymbol{E_r}|}{AF(\phi, \theta)} \exp\left(\frac{iwP_L}{c}\right) \tag{5.23}$$

where V_r = received voltage, $AF(\phi, \theta)$ = Antenna factor, $PLF(\phi, \theta)$ = Polarisation loss factor, $\boldsymbol{E_r}$ = magnitude of Electric field for a path length, P_L, w = frequency

phasor, $2\pi f$, and c = speed of light. As before, the complex received voltages then have bandpass Gaussian noise samples of noise figure of 4 dB are added before the modulus is taken so that signal strength estimates can be processed by the adjacent strongest receiver algorithm. The signal strengths from the two strongest adjacent receivers are then used to calculate a measured bearing using a lookup table containing the logarithmic voltage ratio $20\log_{10}(\tan\phi)$ versus azimuthal angle of arrival. The measured bearing is then compared against the actual bearing by first converting both bearings into complex form and then multiplying the actual bearing by the complex conjugate of the measured bearing. The net result is the angular bearing error measured by the reader. This is recorded and the process continues by examining the next transmitter position until the transmitter grid is complete.

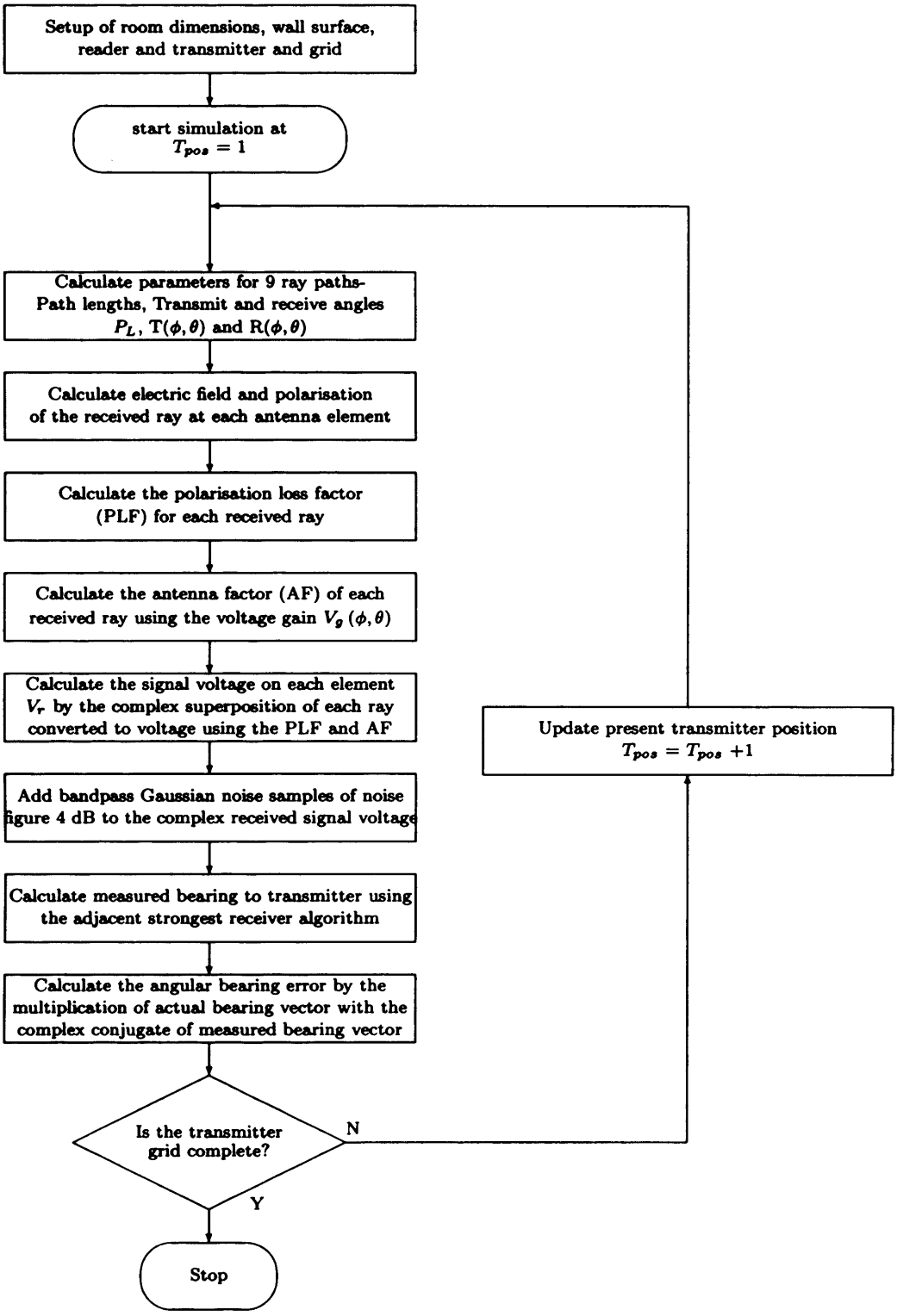
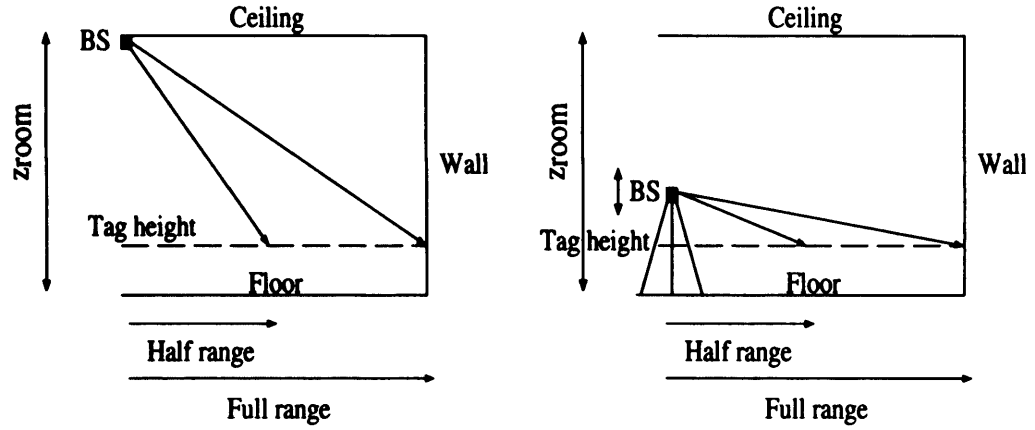


Figure 5.12: Program sequence for model B

The simulation also provides a mechanism for the adjustment of receiver element tilt in elevation. This feature is particularly important for readers mounted at a greater height than the transmitter perhaps on a tripod as shown in Figure 5.13(b) or even attached to the ceiling as shown in Figure 5.13(a).



(a) Ceiling mounting

(b) Tripod mounting

Figure 5.13: A comparison of receiver antenna element pointing angle for a ceiling reader mounting and a tripod reader mounting

The antenna elements are pointed at an appropriate tilt angle so as to maximise the amount of signal captured whether from LOS or NLOS tag transmissions. If the antenna elements are pointed at a tilt angle corresponding to the maximum horizontal range the received signals from this direction would be maximised whilst creating a deadzone for signals received from the regions from mid-range to directly beneath the base station. The optimum element tilt angle for highest probability of reception, assuming an unknown user movement within the room and a base station placed at ceiling level in the centre of the room, should be halfway between zero horizontal range and the maximum horizontal range.

Comparing the pointing angles for reader placed on a ceiling at 3.5 m height and a reader placed on a tripod at 1.8 m height. For both scenarios the transmitter is worn at a height of say 1.2 m. If we first consider the room to be square and of size 25 x 25 m then the maximum horizontal range is 12.5 m. The tilt angle is calculated

using the horizontal pointing range of $\frac{1}{4}$ of the room width. The tilt angle is equal to the inverse tangent of the vertical distance between the tag and the receive element divided by the mid range distance given as

$$\phi_{\text{rtilt}} = \tan^{-1}\left(\frac{\text{reader height} - \text{transmitter height}}{\text{mid range distance}}\right) \tag{5.24}$$

The reader mounted on the ceiling has ϕ_{rtilt} equal to 20.02° and the reader mounted on the tripod has ϕ_{rtilt} equal to 5.48°. It is apparent that a lower tripod height would provide better angular bearing accuracy than the higher ceiling mounting because the default logarithmic power ratio $20\log_{10}(\tan\phi)$ function at zero elevation is used in the calculation of the angle of arrival since the location system can not determine changes in elevation from the measurements taken. An added benefit of the tripod mounting is a much reduced deadzone which reduces the probability of not identifying all the RF-ID throughout the reading space.

5.3.3 Results and discussion on the angular bearing accuracy from a single reader using model B

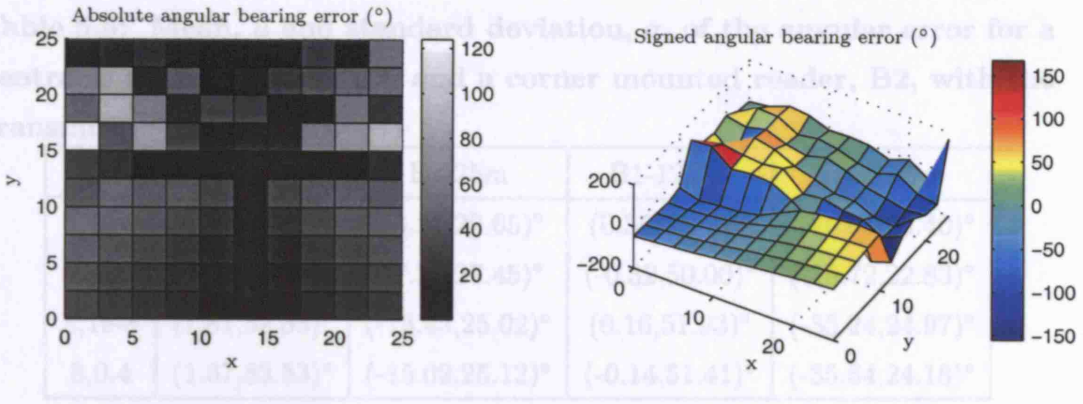
For the results to be presented here, the transmitter was moved in discrete steps of 2.5 m on a grid in x and y around a 25 m square room for a total of 625 transmitter positions. At each transmitter position, the angular bearing error was recorded using model B with a wall surface of relative dielectric constant, ε_r and conductivity, σ, by a stationary reader placed either in the centre (B1) of the room or in the corner (B2) of the room. The reader in the corner requires only two antenna elements to cover the 90° field of view while the central mounted reader requires four antenna elements to cover the 360° field of view. The room size, wall conductivity, σ, wall relative dielectric constant, ε_r, were all varied and the measurements of angular bearing for the grid of transmitter positions were repeated. Specifically, the room sizes examined by the simulation were 25 m square and 12.5 m square with the grid step size scaled for the smaller room size to maintain 625 transmitter positions. Similarly, the reflecting surface relative dielectric constant, ε_r and the conductivity, σ, were examined for four different values, two of which were described in Figure 5.2, $\varepsilon_r = 4, \sigma = 1e-8$, $\varepsilon_r = 4, \sigma = 0.4$, $\varepsilon_r = 8, \sigma = 1e-8$ and $\varepsilon_r = 8, \sigma = 0.4$.

Likewise the model was simulated for five transmitter pointing angles 0°, 90°, 180°, −90° and 135° in azimuth for each position location because this ensured a fair evaluation of the angular bearing error when using a transmitter antenna element with a directional field pattern. The statistics of the angular bearing errors are listed in Tables (5.5, 5.6, 5.7, 5.8, 5.9) for a reader positioned at the same height as the transmitter and containing receive antenna elements with zero tilt in elevation. The angular bearings are also shown using a series of plots of the angular bearing errors for each pointing angle when the simulation was performed on a room of size 25 m and the reader was placed both in the bottom left hand corner of the room and in the centre of the room. The plots for each pointing angle compare the absolute errors for each position with an overhead grayscale view against the signed errors using a colourmap with view from 30° azimuth and 60° elevation.

The tabulated angular errors shown on the preceeding pages for each of the five transmitter pointing angles show that the angular error for the corner reader mounting and central reader mounting can vary significantly with the pointing angle of the transmitter. For example for a 25 m square room the angular accuracy of the corner reader mounting varies from a mean of −0.013° and a standard deviation of 0.23° at a transmitter pointing angle of 135° as seen in Table 5.9 to a mean error of 38.94° and a standard deviation of 21.45° at a pointing angle of 90° as seen in Table 5.6. Similarly the angular accuracy of the central reader mounting varies from a mean of −0.71° and standard deviation of 52.51° at a transmitter pointing angle of 135° as seen in Table 5.9 to a mean angular error of 2.57° and a standard deviation of 68.63° at a pointing angle of 180° as seen in Table 5.7. The variation in the angular error with mounting and transmitter pointing angles are broadly similar for the four different dielectric constants and for the two room sizes.

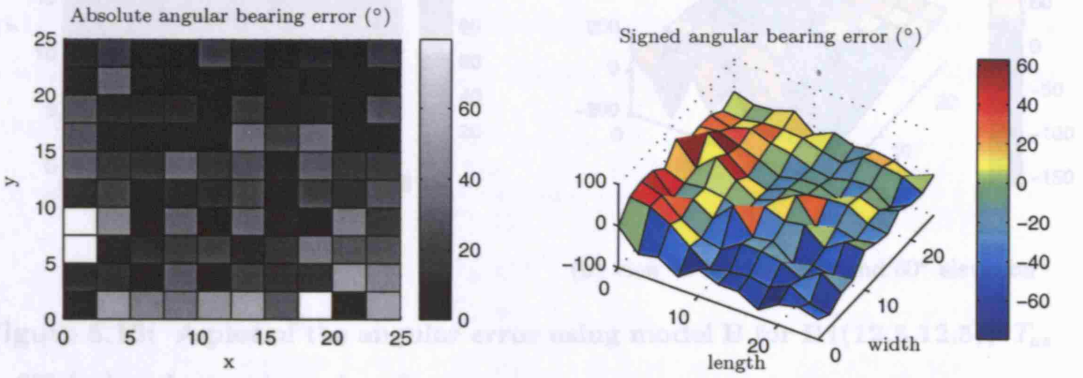
We can therefore expect the angular accuracy of the corner mounted reader to have very low bias, close to zero, and low variance, less than a degree, when at least two of the receive elements on the reader are within the beamwidth of the transmit antenna. The directional field pattern of the antenna on the transmitter not only provides a line of sight signal but also limits the number and strength of any

multipath signals. However when the transmitter pointing angle is not favourable the angular accuracy can become very variable with the potential for large angular errors, the mean and standard deviation of the angular error could both be in excess 30°. It is however more likely that the angular error will be somewhere in between these two extremes such as found in the angular errors corresponding to the transmitter pointing angle of 90° as seen in Table 5.8 where the mean angular error is around 8 - 15 ° and the standard deviation is around 25 - 29 °. In contrast we can expect the central mounted reader to provide an angular error with mean error, no more than 8 °, regardless of the direction the transmitter is pointing but with very high variance, the standard deviation of the angular error was found to be at least 50° when compared to the largest variance found with the corner mounted reader. If we also take into account the fact that the angular bearing error for the central mounted reader can lie anywhere in the range ± 180° whilst the angular bearing error for the corner mounted reader is restricted to the range ± 90° the central mounted reader is more likely to show larger variations in the measurement of the angle of arrival to the transmitter. It is therefore apparent that on balance the corner mounted reader is more likely to provide better accuracy than a central mounted reader and as a consequence the mounting of the corner readers requiring only two antenna elements represents a more judicious choice of reader placement.



(a) plan view (b) view at 30° azimuth and 60° elevation

Figure 5.14: A plot of the angular error using model B for B1(12.5,12.5), T_{az} = 0° (↑) and $\varepsilon_r = 4, \sigma = 1e-8$

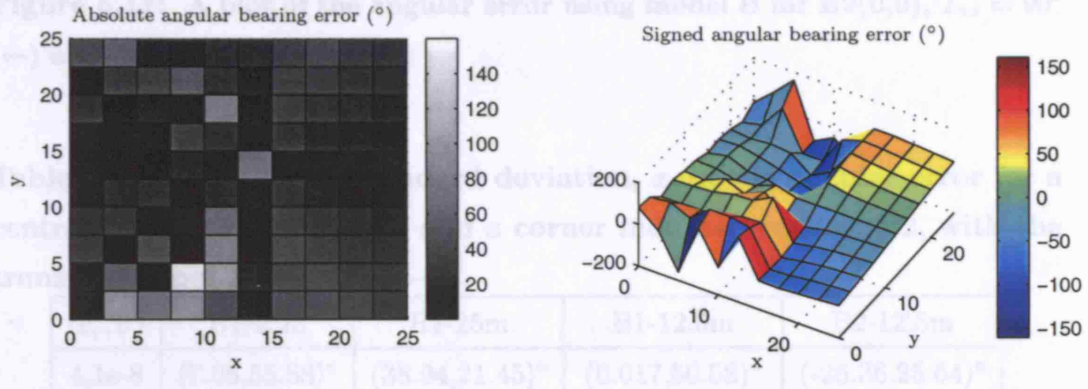


(a) plan view (b) view at 30° azimuth and 60° elevation

Figure 5.15: A plot of the angular error using model B for B2(0,0), $T_{az} = 0°$ (↑) and $\varepsilon_r = 4, \sigma = 1e-8$

Table 5.5: Mean, μ and standard deviation, σ, of the angular error for a central mounted reader, B1, and a corner mounted reader, B2, with the transmitter pointed at 0° (↑)

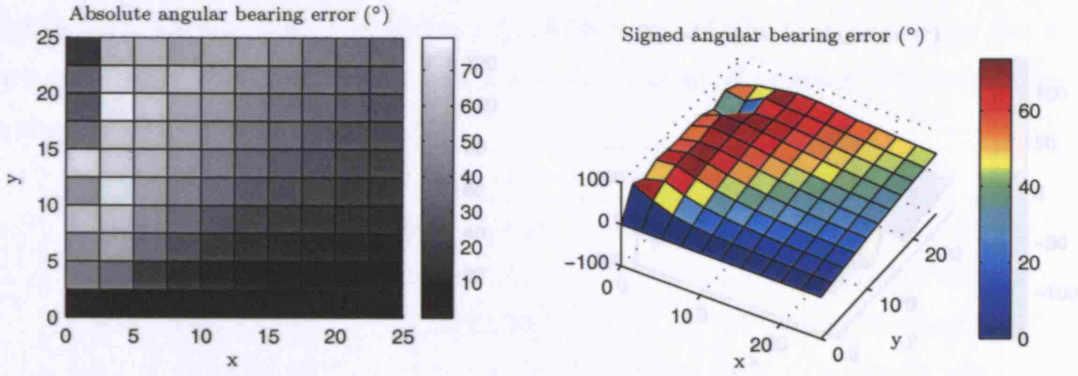
(ε_r, σ)	B1-25m	B2-25m	B1-12.5m	B2-12.5m
4,1e-8	(0.38,53.01)°	(-3.16,29.65)°	(0.35,50.46)°	(-26.08,25.40)°
4,0.4	(0.50,56.21)°	(-7.56,27.45)°	(-0.32,50.00)°	(-25.72,22.83)°
8,1e-8	(1.81,52.33)°	(-13.45,25.02)°	(0.16,51.93)°	(-35.24,24.97)°
8,0.4	(1.67,53.53)°	(-15.02,25.12)°	(-0.14,51.41)°	(-35.54,24.18)°



(a) plan view

(b) view at 30° azimuth and 60° elevation

Figure 5.16: A plot of the angular error using model B for B1(12.5,12.5), T_{az} = 90° (←) and $\varepsilon_r = 4, \sigma = 1e-8$



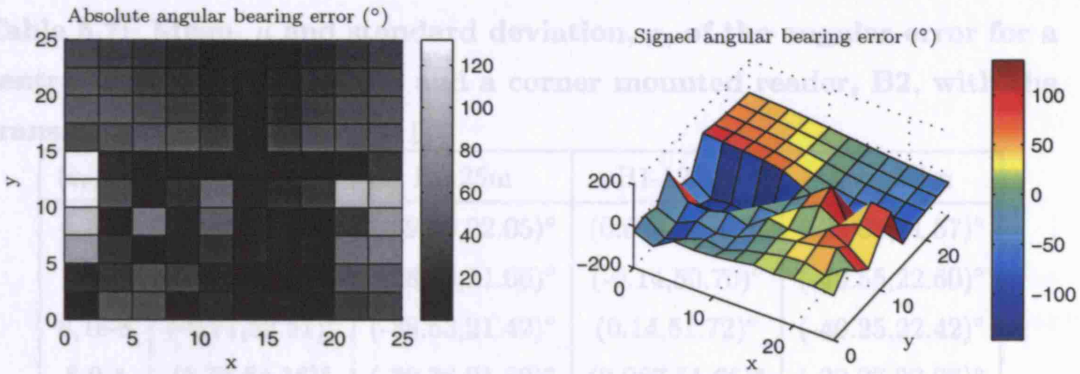
(a) plan view

(b) view at 30° azimuth and 60° elevation

Figure 5.17: A plot of the angular error using model B for B2(0,0), $T_{az} = 90°$ (←) and $\varepsilon_r = 4, \sigma = 1e-8$

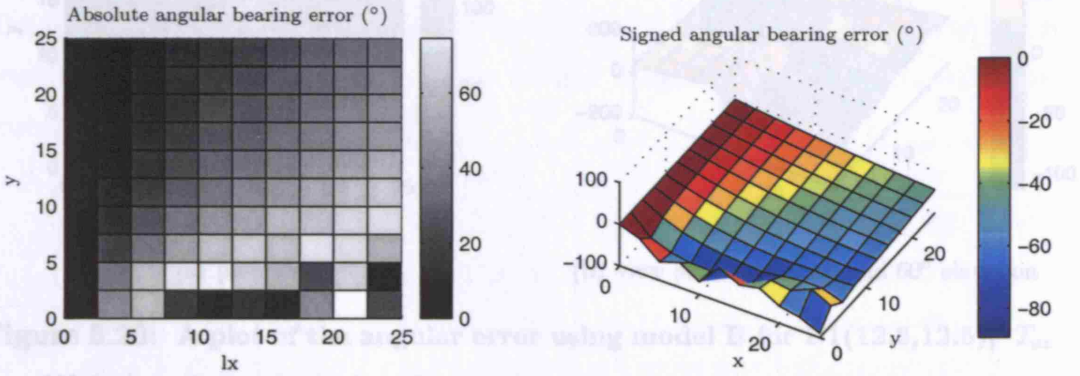
Table 5.6: Mean, μ and standard deviation, σ, of the angular error for a central mounted reader, B1, and a corner mounted reader, B2, with the transmitter pointed at 90° (←)

(ε_r, σ)	B1-25m	B2-25m	B1-12.5m	B2-12.5m
4,1e-8	(7.05,55.88)°	(38.94,21.45)°	(0.017,50.58)°	(-26.36,25.64)°
4,0.4	(1.50,55.14)°	(37.82,21.53)°	(-0.052,51.94)°	(39.83,22.06)°
8,1e-8	(-2.0,53.84)°	(38.77,21.67)°	(-0.72,51.64)°	(40.89,22.77)°
8,0.4	(1.13,53.70)°	(39.31,21.47)°	(0.07,51.61)°	(40.57,22.64)°



(a) plan view (b) view at 30° azimuth and 60° elevation

Figure 5.18: A plot of the angular error using model B for B1(12.5,12.5), T_{az} = 180° (↓) and $\varepsilon_r = 4, \sigma = 1e-8$

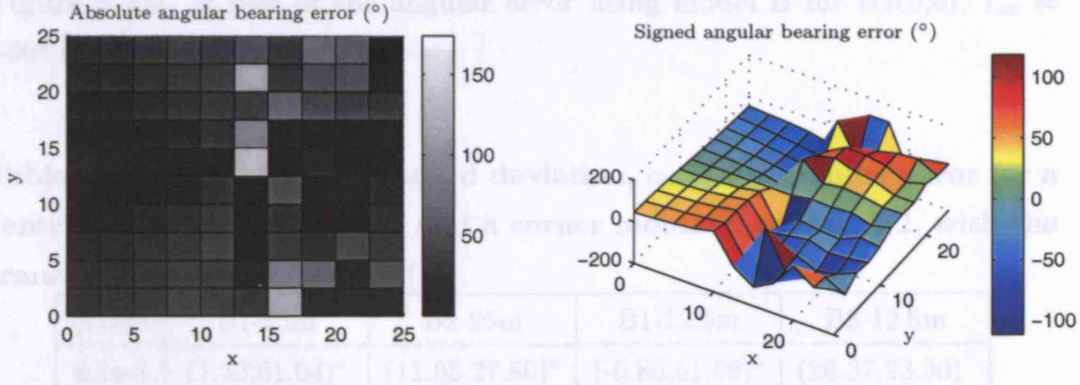


(a) plan view (b) view at 30° azimuth and 60° elevation

Figure 5.19: A plot of the angular error using model B for B2(0,0), T_{az} = 180° (↓) and $\varepsilon_r = 4, \sigma = 1e-8$

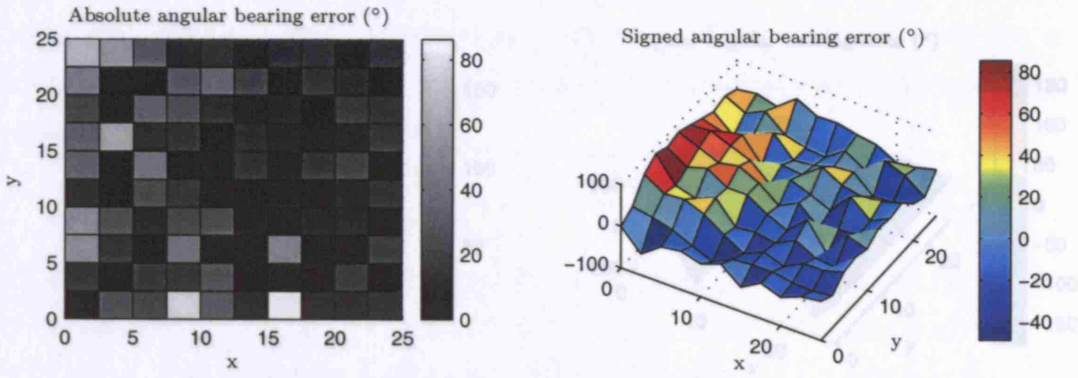
Table 5.7: Mean, μ and standard deviation, σ, of the angular error for a central mounted reader, B1, and a corner mounted reader, B2, with the transmitter pointed at 180° (↓)

(ε_r, σ)	B1-25m	B2-25m	B1-12.5m	B2-12.5m
4,1e-8	(1.99,57.04)°	(-39.89,22.05)°	(0.093,52.43)°	(-39.03,21.57)°
4,0.4	(-3.89,55.33)°	(-38.90,21.66)°	(-0.14,50.70)°	(-40.55,22.60)°
8,1e-8	(-0.71,52.51)°	(-38.53,21.42)°	(0.14,51.72)°	(-40.25,22.42)°
8,0.4	(1.73,54.16)°	(-39.36,21.82)°	(0.067,51.66)°	(-39.27,22.06)°



(a) plan view (b) view at 30° azimuth and 60° elevation

Figure 5.20: A plot of the angular error using model B for B1(12.5,12.5), T_{az} = −90° (→) and $\varepsilon_r = 4, \sigma = 1e-8$



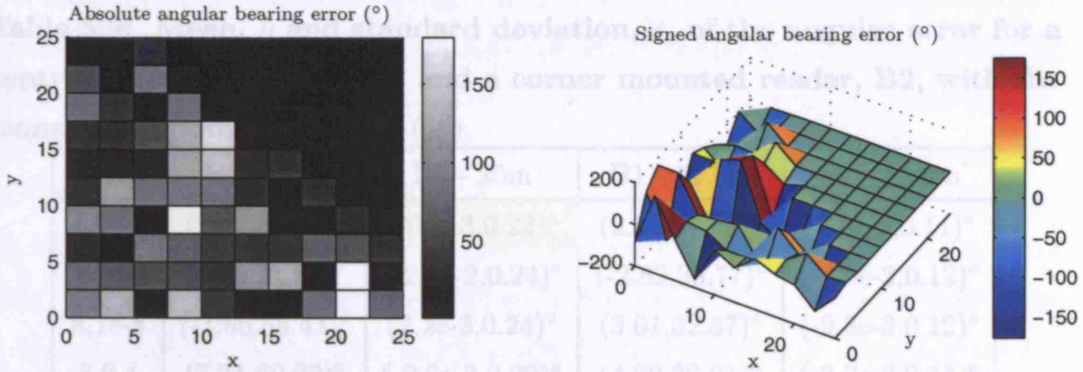
(a) plan view

(b) view at 30° azimuth and 60° elevation

Figure 5.21: A plot of the angular error using model B for B2(0,0), $T_{az} = -90°$ (→) and $\varepsilon_r = 4, \sigma = 1e-8$

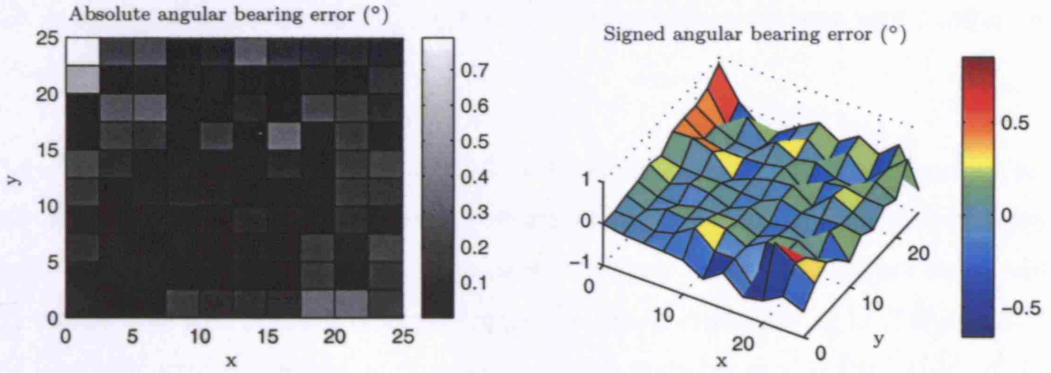
Table 5.8: Mean, μ and standard deviation, σ, of the angular error for a central mounted reader, B1, and a corner mounted reader, B2, with the transmitter pointed at $-90°$ (→)

(ε_r, σ)	B1-25m	B2-25m	B1-12.5m	B2-12.5m
4,1e-8	(1.33,61.04)°	(11.05,27.80)°	(-0.80,51.69)°	(26.37,23.30)°
4,0.4	(-4.74,51.46)°	(8.20,29.93)°	(-0.27,52.23)°	(25.90,24.21)°
8,1e-8	(0.75,51.12)°	(15.06,25.19)°	(-0.11,51.95)°	(33.94,24.45)°
8,0.4	(2.09,54.44)°	(13.88,25.00)°	(-0.32,50.73)°	(34.20,24.11)°



(a) plan view (b) view at 30° azimuth and 60° elevation

Figure 5.22: A plot of the angular error using model B for B1(12.5,12.5), T_{az} = 135° (↙) and $\varepsilon_r = 4, \sigma = 1e-8$



(a) plan view (b) view at 30° azimuth and 60° elevation

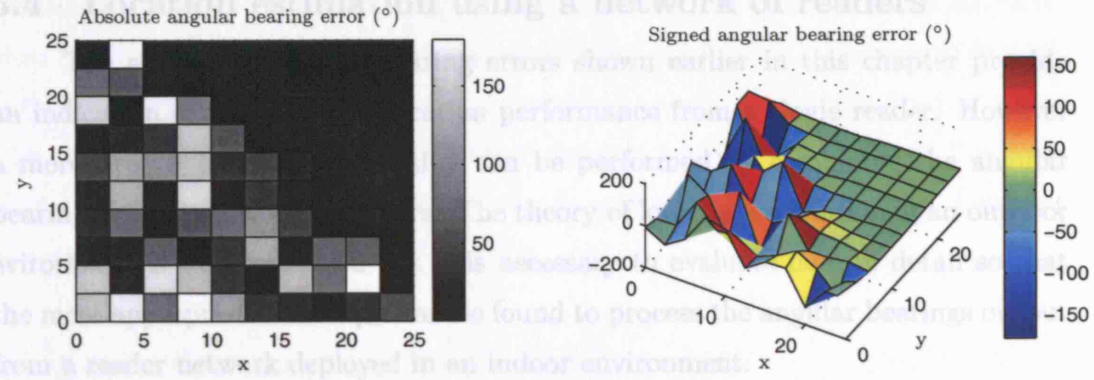
Figure 5.23: A plot of the angular error using model B for B2(0,0), T_{az} = 135° (↙) and $\varepsilon_r = 4, \sigma = 1e-8$

Table 5.9: Mean, μ and standard deviation, σ, of the angular error for a central mounted reader, B1, and a corner mounted reader, B2, with the transmitter pointed at 135° (↙)

(ε_r, σ)	B1 - 25m	B2 - 25m	B1 - 12.5m	B2 - 12.5m
4,1e-8	(2.57,68.63)°	(-0.013,0.23)°	(2.16,37.07)°	(7.3e-2,0.11)°
4,0.4	(1.66,71.91)°	(-2.6e-2,0.24)°	(-2.82,33.77)°	(-4.5e-3,0.12)°
8,1e-8	(-1.56,58.43)°	(2.2e-3,0.24)°	(3.01,32.37)°	(-9.8e-3,0.12)°
8,0.4	(7.71,60.00)°	(-2.9e-3,0.22)°	(4.09,32.91)°	(-3.7e-3,0.11)°

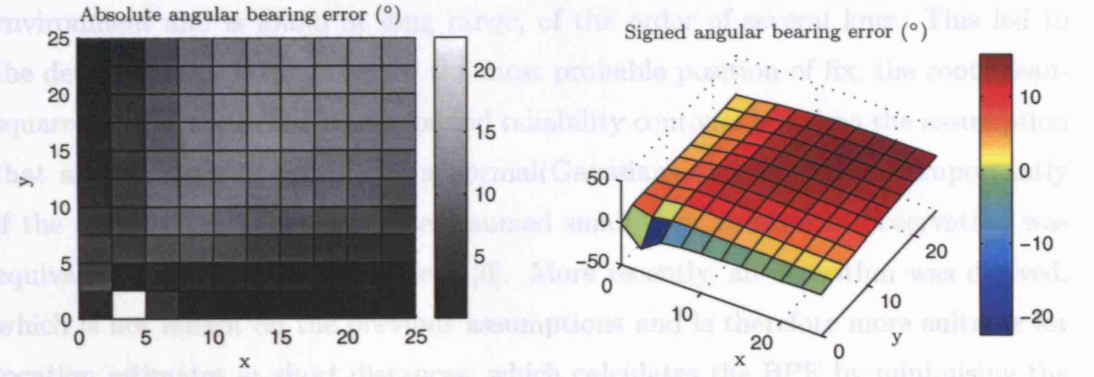
The accuracy of the reader mountings are now reconsidered whenever elevation uncertainty is added into the estimation of the angle of arrival from direction array finding. The receiver elements are tilted by an angle of 11° and the simulation repeated for a transmitter pointing angle of 135° for both the central mounted reader and the corner mounted reader. The receiver tilt is introduced using a reader modelled at a height of 1.8 m and the transmitter modelled at a height of 1.2 m. The plots of angular error for a 25 m square room are shown for both reader mountings in Figures 5.24 and 5.25. The corner mounting now provides a mean angular error of 8.96° and a standard deviation of 6.06°. Whereas the central mounting now provides a mean value of 3.36° and a standard deviation of 65.86°. These values are significantly larger than those highlighted in Table 5.9. This finding suggests that limiting elevation ambiguity is particularly important for this AOA technique because the logarithmic voltage ratio $20\log_{10}(\tan\phi)$ between adjacent receiver channels uniquely defines the angle of arrival in azimuth only without any corrections for elevation differences.

The results from model B have demonstrated that reader locations and elevation are only secondary causes of angular error, above all else the availability of a LOS signal to the reader has the greatest effect on accuracy since it is the strongest signal and determines the true bearing to the transmitter. The availability of LOS signal from the transmitter to the reader is dependent on the pointing angle of the transmitter because the radiation field from the tag antenna is directional and does not radiate much beyond ±90°. Consequently the angular bearing estimates are more at risk of being unreliable.



(a) plan view (b) view at 30° az and 60°el

Figure 5.24: A plot of the angular error using model B for B1(12.5,12.5), T_{az} = 135° (↙) $R_{tilt} = 11°$ and $\varepsilon_r = 4, \sigma = 1e-8$



(a) plan view (b) view at 30° azimuth and 60° elevation

Figure 5.25: A plot of the angular error using model B for B2(0,0), T_{az} = 135° (↙) $R_{tilt} = 11°$ and $\varepsilon_r = 4, \sigma = 1e-8$

5.4 Location estimation using a network of readers

The estimated angular bearing errors shown earlier in this chapter provide an indication of the expected location performance from a single reader. However a more precise estimate of location can be performed by combining the angular bearings from a network of readers. The theory of location estimation in an outdoor evironment is well establised but it is necessary to evaluate here in detail so that the most appropriate technique can be found to process the angular bearings output from a reader network deployed in an indoor environment.

Location estimation involves first finding the most probable location of an object from a set of bearings, referred to in the literature as the BPE (best point estimate) and second finding out the likely error in this estimate. This problem has been examined most often in the context that the target to be located is in an outdoor environment and is found at long range, of the order of several kms. This led to the development of equations for the most probable position of fix, the root-mean-square error in the position of a fix and reliability contours based on the assumption that angular bearing errors had a normal(Gaussian) distribution and importantly if the angular bearing errors were assumed small then an error of observation was equivalent to a lateral displacement [6]. More recently, an algorithm was derived, which is not reliant on the previous assumptions and is therefore more suitable for location estimates at short distances, which calculates the BPE by minimising the the square of the distance between from the measured lines of bearing [36]. This algorithm is outlined in detail in the following section.

5.4.1 Location estimation theory

The calculation of the BPE (best point estimate) is based on minimizing the square of the error distance of the BPE from the measured line of positions as shown in Figure 5.26. Before defining the error distance, d_i, we must first specify the various parameters of the network. The base stations are at co-ordinates (x_i,y_i) and the bearing from the i^{th} system is ϕ_i. The BPE has co-ordinates calculated as (x_t,y_t) and the bearing error for the i^{th} system is $\Delta\phi_i$. Having previously specified

the various parameters, the error distance, d_i, is computed using the perpendiculars from the BPE to the lines of position from each base station. The length of each perpendicular can be deduced using the geometry of the lines as derived in Equations 5.25 to 5.28. This provides a convenient analytic solution which would not have been possible if the miss distance was chosen using either the horizontal distances or vertical distance to the lines of position.

$$H \Rightarrow \frac{x_t - x_i}{\cos(\phi_i)} \tag{5.25}$$

$$O = H\sin(\phi_i) \Rightarrow \frac{(x_t - x_i)\sin(\phi_i)}{\cos(\phi_i)} \tag{5.26}$$

$$h = O - (y_t - y_i) \Rightarrow \frac{(x_t - x_i)\sin(\phi_i)}{\cos(\phi_i)} - y_t + y_i \tag{5.27}$$

$$d_i = h\cos(\phi_i) \quad = (x_t - x_i)\sin(\phi_i) - y_t\cos(\phi_i) + y_i\cos(\phi_i) \tag{5.28}$$

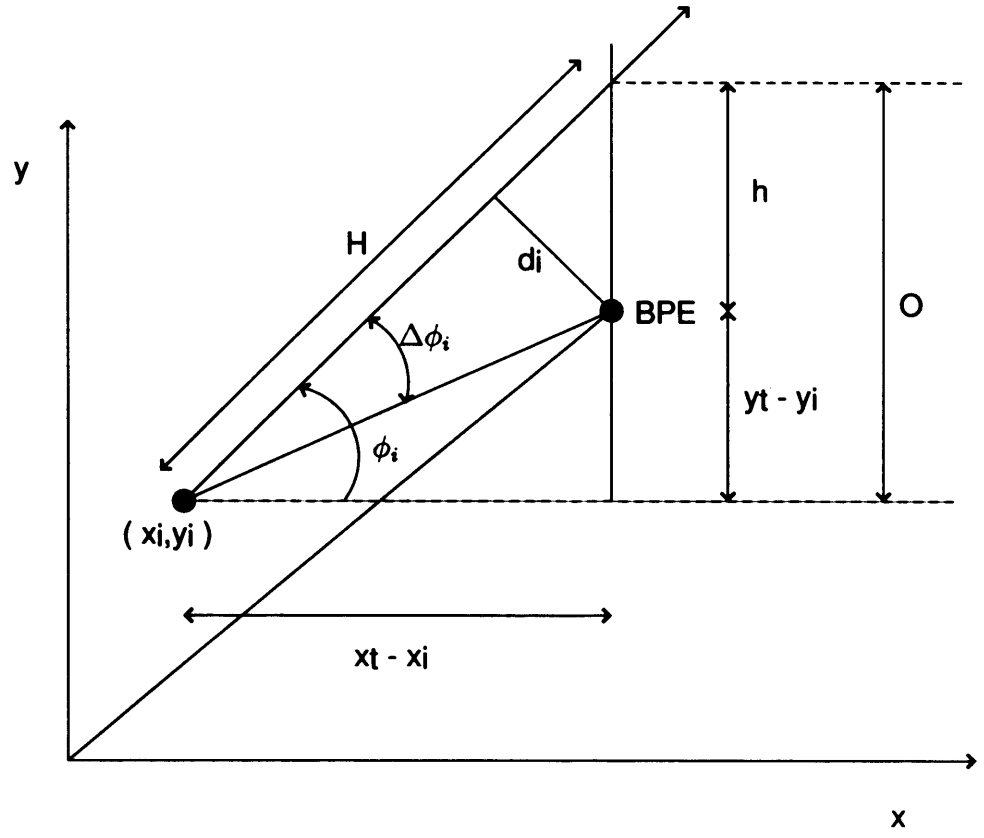


Figure 5.26: An illustration of the calculation of BPE in two dimensions using an algorithm based on minimizing the square of the error distance of the BPE from the measured line of positions

The next step in the calculation of the BPE is to choose the function to be minimised. Minimising the function d_i is not appropriate because the use of the absolute value results in discontinuous derivatives which cannot be treated analytically. A more appropriate function to minimize is the sum of squares of the total miss distance. The function d_i^2 must be differentiated with respect to all i systems, using the constituent error distances from each system. The differentiation of the function d_i^2 does not minimize the actual deviations from each line which are be measured perpendicular to the given function, rather it de-emphasizes large outliers (negative and positive results are not able to cancel each other out). The resulting

residual, D, shown in Equation 5.29 is then differentiated to find the BPE.

$$\begin{aligned} D &= \sum_{i=1}^{N} a_i^2 x_t^2 + \sum_{i=1}^{N} 2a_i b_i x_t y_t - \sum_{i=1}^{N} 2a_i c_i x_t \\ &\quad + \sum_{i=1}^{N} b_i^2 y_t^2 - \sum_{i=1}^{N} 2b_i c_i y_t + \sum_{i=1}^{N} c_i^2 \end{aligned} \tag{5.29}$$

where $a_i = \sin\phi_i$; $b_i = -\cos\phi_i$; $c_i = x_i \sin\phi_i - y_i \cos\phi_i$; N = number of LOPs

The differentiation of D is performed by setting the partial derivatives of D with respect to both x_t and y_t to zero. This will provide the values of x_t and y_t for which the total squared distance is minimised.

$$\frac{\delta D}{\delta x_t} = 2x_t \sum_{i=1}^{N} a_i^2 + 2y_t \sum_{i=1}^{N} a_i b_i - 2\sum_{i=1}^{N} a_i c_i = 0 \tag{5.30}$$

$$\frac{\delta D}{\delta y_t} = 2x_t \sum_{i=1}^{N} a_i b_i + 2y_t \sum_{i=1}^{N} b_i^2 - 2\sum_{i=1}^{N} b_i c_i = 0 \tag{5.31}$$

The values of x_{tmin} and y_{tmin} which minimise the derivatives can then be then calculated by rearranging the relationship

$$\frac{\delta D}{\delta x_t} = \frac{\delta D}{\delta y_t}$$

as a result the values can be expressed as

$$x_{tmin} = \frac{\sum_{i=1}^{N} a_i c_i - y_t \sum_{i=1}^{N} a_i b_i}{\sum_{i=1}^{N} a_i^2} \tag{5.32}$$

$$y_{tmin} = \frac{\sum_{i=1}^{N} b_i c_i - x_t \sum_{i=1}^{N} a_i b_i}{\sum_{i=1}^{N} b_i^2} \tag{5.33}$$

The calculation of BPE - (x_t, y_t) - is completed by the cross substitution of the x_{tmin} and y_{tmin}.

$$x_t = \frac{\sum_{i=1}^{N} b_i^2 \sum_{i=1}^{N} a_i c_i - \sum_{i=1}^{N} a_i b_i \sum_{i=1}^{N} b_i c_i}{\sum_{i=1}^{N} a_i^2 \sum_{i=1}^{N} b_i^2 - \left(\sum_{i=1}^{N} a_i b_i\right)^2} \tag{5.34}$$

$$y_t = \frac{\sum_{i=1}^{N} a_i^2 \sum_{i=1}^{N} b_i c_i - \sum_{i=1}^{N} a_i b_i \sum_{i=1}^{N} a_i c_i}{\sum_{i=1}^{N} a_i^2 \sum_{i=1}^{N} b_i^2 - \left(\sum_{i=1}^{N} a_i b_i\right)^2} \tag{5.35}$$

The most probable position fix is estimated by Equations 5.34 and 5.35. While this provides a position fix it does not indicate the likelihood that this position is the true position. An alternative approach that provides both the location estimate and the associated probability spread is to minimise the error distance d_i and then calculate the least square estimate and covariance from this function.

If we reformulate Equation 5.28 using the constants as previously chosen, then the error distance $D = a_i x_t + b_i y_t - c_i$. The error distance, D, as defined in 5.36 consists of a set of linear equations. As a result a least square error estimator can then be used to estimate the location provided there is a minimum of three base stations with 3 lines of position.

$$\mathbf{D} = \mathbf{HP} - \mathbf{C} \tag{5.36}$$

where:

$$C = \begin{bmatrix} c_o \\ c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} \quad P = \begin{bmatrix} x_t \\ y_t \end{bmatrix} \quad H = \begin{bmatrix} a_0 & b_0 \\ a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_N & b_N \end{bmatrix} \quad D = \begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix} \tag{5.37}$$

The computation of the best point estimate involves minimizing the Euclidean norm squared of the residual $\mathbf{HP} - \mathbf{C}$. Using the fact that the squared norm of $\mathbf{v}$ is $\mathbf{v}^T\mathbf{v}$

and by adding a weight matrix $\mathbf{W}$ we can rewrite the expression as

$$\begin{aligned} S(\mathbf{P}) &= (\mathbf{HP}-\mathbf{C})^T\mathbf{W}(\mathbf{HP}-\mathbf{C}) \\ &= (\mathbf{HP})^T\mathbf{W}(\mathbf{HP})-\mathbf{C}^T\mathbf{WHP}-(\mathbf{HP})^T\mathbf{WC}+\mathbf{C}^T\mathbf{WC} \end{aligned} \quad (5.38)$$

The two middle terms are equal, and the minimum is found at the zero of the derivative with respect to $\mathbf{P}$.

$$2\mathbf{H}^T\mathbf{WHP}-2\mathbf{H}^T\mathbf{WC}=0 \quad (5.39)$$

$$\mathbf{H}^T\mathbf{WHP}=\mathbf{H}^T\mathbf{WC} \quad (5.40)$$

Therefore the minimizing vector $\mathbf{P}$ is a solution of the normal equation and can be given as

$$\hat{\mathbf{P}}=\left(\mathbf{H}^T\mathbf{WH}\right)^{-1}\mathbf{H}^T\mathbf{WC} \quad (5.41)$$

The solution can be calculated provided the matrix $\mathbf{H}^T\mathbf{WH}$ has full rank and is well conditioned. The variance of the estimator is the covariance matrix, $\mathbf{Q}$, of the estimate $\hat{\mathbf{P}}$. By error propagation it is calculated as the inverse of the second derivative of $S(\mathbf{P})$ and can be shown as

$$\mathbf{Q}=\left[\mathbf{H}^T\mathbf{WH}\right]^{-1}=\begin{bmatrix}\sigma_x^2 & \rho\sigma_x\sigma_y\\ \rho\sigma_x\sigma_y & \sigma_y^2\end{bmatrix} \quad (5.42)$$

The solution position $\hat{\mathbf{P}}$ and covariance matrix $\mathbf{Q}$ are calculated by a few matrix operations with no iterations. Of course it is not essential to have absolute weightings as relative weighting determines the position of the BPE. However absolute weights are required by the covariance matrix in order to get an accurate understanding of location error spread and related probabilities. The accuracy of the estimator can be further improved by adjusting the values of the weighting matrix inversely in relation to the distance of the current location estimate from each of the bases. Since faraway bases must be weighed less than closer ones because for a fixed error in angle measurement the uncertainty in position increases with distance. The

weighting can also be adjusted inversely to the variance of the angular bearing of each reader provided this is known a priori.

5.4.2 Characterisation of location errors

If we assume a Gaussian distribution for the angular bearing errors at each reader the pdf of the position vector, $\mathbf{P}$, can then be characterised in the x,y plane by elliptical contours of constant probability. These curves assign a probability for a certain ellipsoid size. The locus of the ellipse is the position estimate, $\hat{\mathbf{P}}$, with the covariance matrix, $\mathbf{Q}$, defining the amount of uncertainty and the orientation of the ellipse axis in the x,y plane. The probability density function is defined in Equation 5.43 with the exponent described by the sum of squared deviations and the inverse of the covariance matrix.

$$p(\mathbf{P}) = \frac{1}{2\pi \left|\mathbf{Q}\right|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{P}-\hat{\mathbf{P}})\mathbf{Q}^{-1}(\mathbf{P}-\hat{\mathbf{P}})} \tag{5.43}$$

A number of points can be made if the exponent is expanded as shown in (5.44). If there is no $P_x P_y$ present, then the principal axes of the ellipse are aligned with the co-ordinate system. The distance from the mean point, $\hat{\mathbf{P}}$, to the covariance ellipse does not describe the standard deviation along directions other than along the principal axes.

$$\frac{\left(P_x - \hat{P}_x\right)^2}{\sigma_x^2} - \frac{2\rho\left(P_x - \hat{P}_x\right)\left(P_y - \hat{P}_y\right)}{\sigma_x \sigma_y} + \frac{\left(P_y - \hat{P}_y\right)^2}{\sigma_y^2} = \kappa \tag{5.44}$$

where κ is a constant that determines the size of the 2 dimensional region and ρ is the correlation coefficient. It can then be proven [37] that the probability of location error of minimum area can be set by the following relationship.

$$Pr = 1 - e^{-\frac{1}{2}\kappa} \tag{5.45}$$

Table 5.10 shows the value of the constant, κ, calculated for a number of different probability values. Thus rather than giving an exact value to the location, $\mathbf{P}$, we can say that given, $\hat{\mathbf{P}}$ and $\mathbf{Q}$, then with 90% probability, the true location is within

in an ellipse centred on, $\hat{\mathbf{P}}$, and a perimeter defined according to the scaling factor $\kappa = 4.605$. Moreover, if we lay no emphasis on the shape and the orientation of the error ellipse the position accuracy can be described by the area of the elliptic region. The drawback is that this measure of error tends to zero for slim ellipses even if there exists a large error in the direction of one of the main axes. In the case of a degeneration to a straight line the area would even become zero.

Pr	κ
50%	1.386
60%	1.832
70%	2.408
80%	3.219
90%	4.605

Table 5.10: The relationship of probability to the ellipse scaling factor

A more commonly used and better suited description involves an error circle. The definition is based on the assumption that the overall root-mean-square error can be described as the square root of the variances along the main axes.

$$D_{rms} = \sqrt{\sigma_x^2 + \sigma_y^2} \tag{5.46}$$

The rms error is not directly related to a specific probability. The latter is dependent on numerical integration of an elliptical pdf contour with a circular boundary. Specifically the probability that the true position is inside the area is related to the axial ratio of the ellipse and amounts to 63% in the case of a circle and 68% if the ellipse degenerates to a straight line [38]. While relating probability to a given area provides some confidence in the position of transmitter when the angular bearing errors are distributed within a Gaussian distribution. If the error distribution is unkown or not measureable then an alternative method of assigning probability to location is a distribution free confidence region [39]. These regions are said to exist for the true position of the transmitter for all probability distributions (not necessarily normal) provided the angular errors are independent and have zero median.

The probability that a true position lies inside the closed polygon can be related directly to the number of lines of bearing n. Since the errors are of zero median then all n^2 combinations of signs are equally likely. An infinite plane can be subdivided into $\frac{1}{2}(n^2+n+2)$ polygons, of which 2n extend to finity and the remaining $\frac{1}{2}(n-1)(n-2)$ makes the largest possible closed polygon. The probability that true position lies in one of the open regions is $\frac{2n}{2^n}$. Since the 2n open regions correspond to the same sign combinations, the largest closed polygon is therefore a confidence region with coefficient $P_n = 1 - \frac{n}{2^{n-1}}$. Table 5.11 shows the probability, P_n, calculated for n lines of bearing. The probability coefficent is best illustrated for three lines of bearing as shown in Figure 5.27.

No. of lines n	Probability P_n
3	25.00%
4	50.00%
5	68.75%
6	81.25%
7	89.06%
8	93.75%

Table 5.11: Probability for n lines of position for a distribution free confidence region

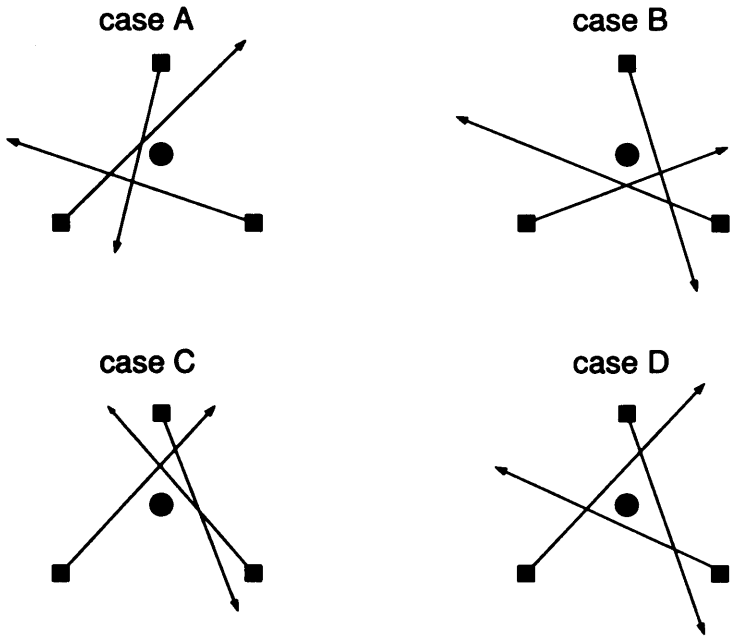


Figure 5.27: Illustrations of the concept of a distribution free confidence region using three readers and a single transmitter

It should be noted that if two or more lines are parallel then the corresponding region is not closed. The bearing sign is chosen so that anti-clockwise movement of the bearing vector from the true position is a positive sign and clockwise movement therefore results in a negative sign. In case D the true position is enclosed by the polygon, sometimes referred to as a cocked hat. The size of the enclosing region and its subsequent area can vary significantly as can be seen in the illustrations. The confidence region therefore does not have optimimum properties. If we encounter a very small cocked hat, the true position will very probably outside it; whereas if we encounter a very large cocked hat, the true position is very likely to be inside. This however is not to say the large cocked hat will provide a more reliable and hence lower error distance to the true position.

Although the probability that the true location of the transmitter is within the cocked hat (enclosed confidence region) formed from the lines of position does not have optimum properties, it does show numerically the relationship between the likelihood that the true transmitter location is within the enclosed region and the number of readers in the network. Thus we have a real measure of the effect of redundancy.

5.4.3 Sources of location error

The location accuracy of the BPE is affected dramatically by wild bearings. These wild bearings are found more often in error distributions that are more peaky than a normal distribution. The presence of additional outliers can be defined using the kurtosis, K, of the error distribution. Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. That is, data sets with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. Data sets with low kurtosis tend to have a flat top near the mean rather than a sharp peak. A uniform distribution would be the extreme case. The kurtosis is defined using the higher order moments of the error distribution and is given as

$$K = \frac{u_4}{u_2^2} - 3 \tag{5.47}$$

where $K = 0$ represents a normal distribution, $K > 0$ represents a positive kurtosis – Leptokurtic distribution and $K < 0$ represents a negative kurtosis – platykurtic distribution. The wild bearings which manifest in these distributions occur because of the channel propagation conditions.

The rejection of wild bearings is a key step in reducing the location error of the BPE. A useful parameter used in this process is the dispersion factor which can be given as

$$E_o = \sum \frac{\theta_j^2}{V_j} \tag{5.48}$$

where: θ_j is the angular errror for the jth line of position from the calculated BPE and V_j is the angular variance at base station j.

The identification of a wild bearing requires a minimum of four line of positions. Because with two bearings there can no evidence of inconistency. With three bearings an unusually large cocked hat suggests that at least one is dubious, but it is not possible to say which. Only with four or more is there any real possibility of applying rational rejection of a wild bearing. To remove a wild bearing, one bearing is eliminated at a time and the BPE is recalculated along with the corresponding

dispersion factor. The best combination of LOPs will be the one associated with the smallest dispersion factor and the bearing excluded can be assumed wild. For instance, we can form five separated four bearing fixes from a set of five LOPs by dropping each bearing in turn; if one of these fixes has a significantly smaller dispersion factor than the others, the suspect bearing is the one omitted from this fix.

Besides the negative affect of wild bearings, multipath and non line of sight signal on the location accuracy expected from a reader network, the relative position of a transmitter within a reader network also has a significant impact on location performance. The relative variation of location accuracy by a network containing three readers positioned on the vertices of an equilateral triangle is summarised in Figure 5.28. A number of observations can be made from the layout of the readers in this arrangement provided one assumes all readers receive a direct line of sight to the transmitter in the absence of multipath. First, the uncertainty in the position of a fix increases rapidly with increase in distance from the base stations. Second, a two reader fix is very unreliable near the line joining the two readers i.e. where the readers are co-linear. Third, by obtaining location estimates using more than two readers, the position and reliability of a fix is determined predominantly by the nearer readers, except when the fix is near the line joining these readers. Consequently additional readers, not only remove the ambiguity caused by co-linear readers, they allow additional redundancy. While this example shows the relative performance for a particular transmitter location within the reader layout, the absolute overall performance will vary considerably whenever the transmission model incorporates additional complexity. First any layout must also contend with strong multipath components, second the likelihood of a non direct line of sight signal to some or all of the readers and third the height difference between reader and transmitter causes an elevation error in the bearing calculation.

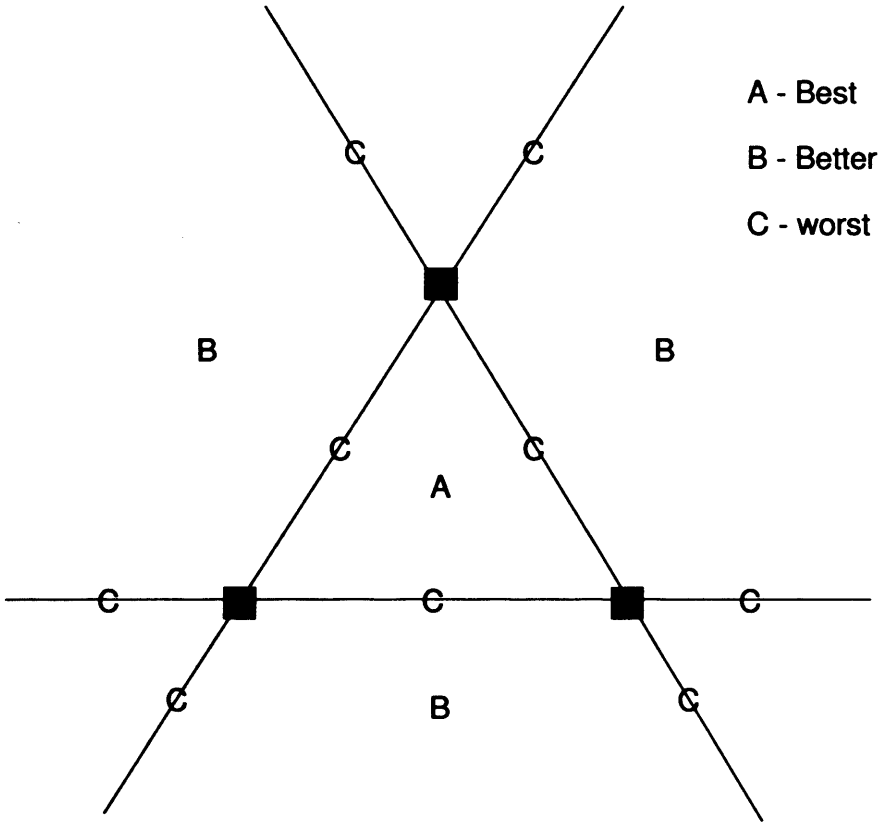


Figure 5.28: Relative variation of location accuracy with respect to position in a three reader direction finding network

While the factors that cause location error are not insurmountable, judicious reader layout is only part of the solution to accurate location estimation. The estimation of location can be improved further by removing wild bearing outliers. Although it should be noted that there is no statistical basis normally for eliminating bearings from the location estimate because in general it is not known which bearings accurately correspond to those from the transmitter and which might not. Ideally making proper use of all available measurement samples to fix the location represents a better solution because it must be emphasised that the location estimate is calculated from a very limited amount information e.g. three instantaneous bearings and by normal statistical standards this must be regarded as a very small number of samples, with the corresponding risk of misleading error. To reduce the risk of misleading location error the angular bearing error at some or all of the readers can potentially be reduced by averaging a series of instantaneous samples each taken at consecutive time instants (equivalent to a snapshot bearing at each reader) assuming the transmission repetition rate can be made much faster than the expected average velocity of the moving person.

5.5 Evaluation of the location accuracy of the reader network topologies using model B

The position error, $\mathbf{P} - \hat{\mathbf{P}}$, from each reader topology is assessed whenever the transmitter moves along an circular path. The accuracy of each reader topology is found using the mean and standard deviation of the position errors recorded. In addition the cumulative probability of the location errors is also recorded to provide a visual comparison of the relative performance for each topology. As well as analysing the position error using the raw bearings everytime the transmitter sends a data burst with repetition interval T_r (equivalent to a repetition rate T_n), the angular bearings are also processed in a group at each reader. This method of processing involves the collection of simultaneous bearing measurements, η, for each reader which are then stored in a buffer. The contents are then processed by a specific algorithm. The algorithm depends on the simultaneity assumption which implies that the bearings in the buffer are collected at the same point. That assumption will not hold in general but to reduce the chances of the assumption being invalid, the size of the buffer is made inversely proportional to the expected speed ie high speed means a smaller buffer (equivalently less samples to average), slow speed means a larger buffer (equivalently more samples to average). Of course in many cases it may be impossible to know the speed a priori. Nevertheless knowing the average walking speed of a person is 1 m/s, the repetition rate of the transmitter can be adjusted accordingly. For instance choosing a repetition interval of 0.1 s provides 10 bearing updates a second, for each of which the person has only moved 0.1 m. This also means a buffer size of $\eta \leq 10$ satisfies the simultaneity conditon. If these samples are used appropriately in conjunction with additional signal processing a reduction in the location error is possible.

5.5.1 Presentation of three distinct averaging algorithms

The algorithms used to process the angular bearings are now outlined. Algorithm I forms an estimate of the location, $\hat{\mathbf{P}}$, using the raw angular bearing measurements at each repetition interval, T_r. The position error, $\mathbf{P} - \hat{\mathbf{P}}_{\mathbf{I}}$, for each transmitter position is then calculated. Algorithm II calculates a position estimate, $\hat{\mathbf{P}}_{\mathbf{II}}$, periodically

every ηT_r seconds by averaging the angular bearings held in the buffer at each reader. The buffer contains a block of samples, η and the resultant averages are then used to calculate the location estimate, $\hat{\mathbf{P}}_{\mathbf{II}}$. This block by block averaging technique can be described in more detail in the following manner. An array of raw complex angular bearings $[y_1, y_2 \ldots, y_{nn}]$ are converted to a new array of smoothed complex angular bearings. The 'smoothed point' S_k is the average of an odd number of consecutive $2n + 1$ $(n = 1, 2, 3, ..)$ points of the raw complex angular bearings y_{k-n}, y_{k-n+1}, ..., y_{k-1}, y_k, y_{k+1}, ..., y_{k+n-1}, y_{k+n}, ie

$$S_k = \sum_{i=-n}^{i=n} y_{\eta k+i} \Big/ (2n+1)$$

The odd number $2n + 1$ is usually named the filter width and here is equal to the block size, η. As before, the corresponding position errors, $\mathbf{P} - \hat{\mathbf{P}}_{\mathbf{II}}$, are calculated with the present position, $\mathbf{P}$, computed by averaging the true transmitter position samples over the period, ηT_r. Algorithm III calculates a position estimate, $\hat{\mathbf{P}}_{\mathbf{III}}$, after each transmitter repetition interval, T_r. It is similar to Algorithm II in that it uses the average angular bearing from a block of size η to estimate the current bearing at the reader but it does this after each transmitter repetition interval, T_r. This approach uses a rectangular or unweighted sliding average smooth to the data and can be given by

$$S_k = \sum_{i=-n}^{i=n} y_{k+i} \Big/ (2n+1)$$

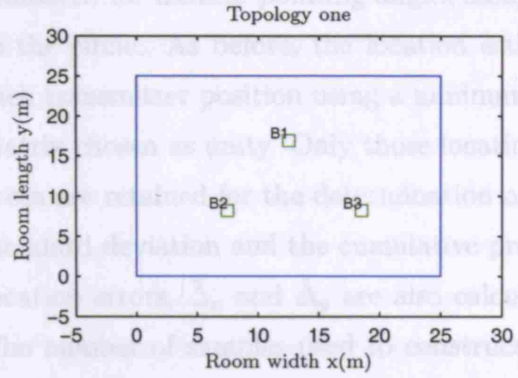
It simply replaces each point in the signal with the average of m adjacent points, where m is a positive integer. For example, for a 3-point smooth (m=3):

$$S_k = \frac{y_{k-1} + y_k + y_{k+1}}{3}$$

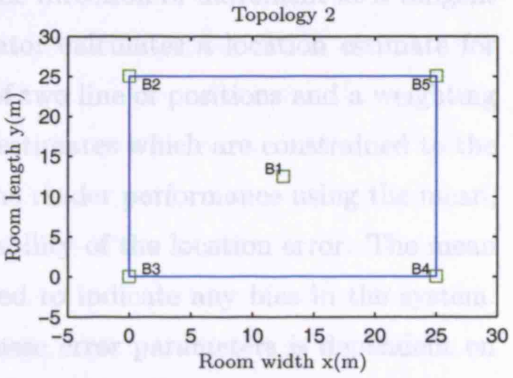
for k=2 to nn - 1, where S_k is the k^{th} point in the smoothed data array, y_k the k^{th} point in the original data array, and nn is the total number of points in the data array. Similar smooth operations can be constructed for any desired smooth width, m. Usually m is an odd number. The smoothing operations can be applied more than once, that is, a previously-smoothed data set can be smoothed again. This

may prove useful but obviously after each filter pass the first n and the last n points are lost.

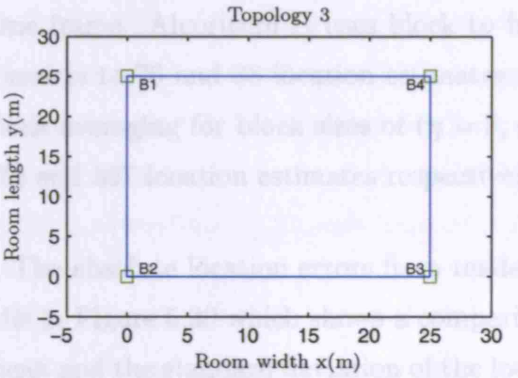
The four reader layouts - topologies one - four - to be analysed for location accuracy are examined in a square room of size 25 m as pictured in Figure 5.29. Topology one contains three readers arranged on the vertices of a centre located isosceles triangle of edge length 10 m. Topology two contains five readers that are arranged in the traditional cellular format with four corner mounted readers and one centrally mounted reader. A further two reader layouts were also examined to verify the performance of network using a lower number of readers. Topology three is similar to topology two by containing four corner mounted readers but with the centrally mounted reader removed. By contrast topology four contains three corner mounted readers that are positioned on a right angled triangle in an L shaped reader layout.



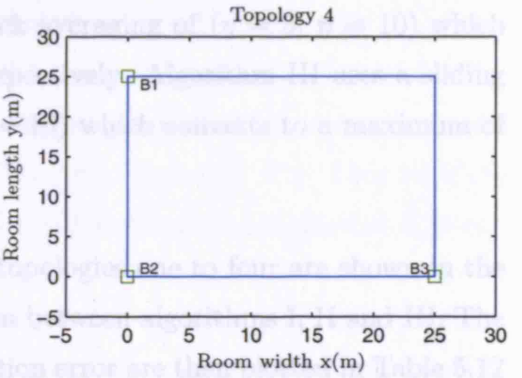
(a) Topology one - three readers arranged in an isoceles triangle with 10 m line distance



(b) Topology two - cellular format with four corner readers and a single central reader



(c) Topology three - square format wirh four corner adjacent readers



(d) Topology four - L format with three corner readers

Figure 5.29: The four topologies of the reader network to be examined for location accuracy

5.5.2 Results and discussion of location accuracy from the reader network topologies

The estimator discussed earlier is applied to the bearings output from reader topologies. The reader topologies are examined using model B for a wall defined as $\varepsilon_r = 4, \sigma = 1e-8$ whenever the transmitter takes moves along a counter-clockwise circular path with centre (12.5,12.5), radius 6 m and start position (18.5,12.5). The transmitter moves with a speed of 1 m/s along the path of the circle and 10 bearings estimates a second are received in a buffer at each reader. The receiver elements on each reader are tilted by an angle of 11° with the reader tripod mounted at a height of 1.8 m and the transmitter worn at a height of 1.2 m. The transmitter is

evaluated for natural pointing angles along the direction of movement at a tangent to the circle. As before, the location estimator calculates a location estimate for each transmitter position using a minimum of two line of positions and a weighting matrix chosen as unity. Only those location estimates which are constrained to the room are retained for the determination of the reader performance using the mean, standard deviation and the cumulative probability of the location error. The mean location errors, $\bar{\Delta}_x$ and $\bar{\Delta}_y$ are also calculated to indicate any bias in the system. The number of samples used to construct these error parameters is dependent on whether or not bearing averaging is applied. Algorithm I applies no averaging and therefore has the largest number of samples at 376 location estimates in a 38 second time frame. Algorithm II uses block to block averaging of ($\eta = 5$, $\eta = 10$) which converts to 76 and 38 location estimates respectively. Algorithm III uses a sliding block averaging for block sizes of ($\eta = 5$, $\eta = 10$) which converts to a maximum of 372 and 367 location estimates respectively.

The absolute location errors from reader topologies one to four are shown in the cdfs in Figure 5.30 which shows a comparison between algorithms I, II and III. The mean and the standard deviation of the location error are then plotted in Table 5.12 to determine the relative performance of each algorithm. Also included in the table for each algorithm is the total number of valid location estimates used to construct the mean and rms values compared to the maximum number of location sample possible from that algorithm. Each location estimate is valid provided the matrix, $\mathbf{H}^T\mathbf{W}\mathbf{H}$, used in the calculation of the position estimates as shown in Equation 5.41, has full rank and the corresponding location estimate is within the dimensions of the room.

We begin the presentation of the location errors by evaluating the performance of algorithm I. The results shown in Table 5.12 for algorithm I indicate that topology two provides the greatest accuracy with mean error of 0.28 m and a standard deviation of 6.96 m. On the other hand topology four provides the lowest accuracy with mean error 1.54 m and a standard deviation of 8.56 m. The results also indicate the averaging algorithms, II and III, are very effective at reducing the presence

of absolute location errors, particularly those in excess of 8 m as seen in the cdfs plotted for topologies two and three as illustrated in Figures 5.30(b) and 5.30(c). The reduction in location error is less pronounced in the cdf plotted for topology four as shown in Figure 5.30(d). Indeed the algorithms have had little or no effect in reducing the location error in the cdf plotted for topology one as shown in Figure 5.30(a). If we now switch our attention to the relative performance of the algorithms. Looking closely at the absolute mean and rms location errors in Table 5.12 it is apparent that the sliding block average algorithm III provides slightly better accuracy than the block by block average algorithm II whether using a block size, η, of ten or five. It also apparent from Table 5.12 that the larger block size, η, of ten also provides marginally better accuracy than a block size of, η, of five for algorithm III. Although the improvement in accuracy from a larger block size is less obvious with algorithm II. As expected the best performing combination of topology and algorithm are topology two and algorithm III with a block size of η of ten with the location errors having a mean error of 0.23 m and a standard deviation of 5.34 m.

The relative performance of the reader topologies can also be compared using the absolute mean and rms location errors of algorithm III using a block size η of ten as tabulated in Table 5.12. Reader topologies two and three have the lowest absolute error after averaging. Topology two that contains five readers has an one more reader than topology three and this extra reader is located at the centre of the room, located 6 m from the points on the transmitter path. Significantly this redundancy provides a lower bias in the mean absolute location error for topology two than topology three. In fact it is apparent that the mean error is close to zero, indicating that there is no bias in the reader network used in topology two. However there is considerable variation around the absolute mean, as much as 5.34 m, which is dependent on the tag location and pointing angle along the circle of movement.

The improvement in location accuracy after averaging is much less visible in reader networks containing three readers such as topology one and four. Topology four that contains three readers, each with two receivers, mounted in adjacent corners of the room has an absolute mean error of 1.44 m and a standard deviation of 6.26 m for

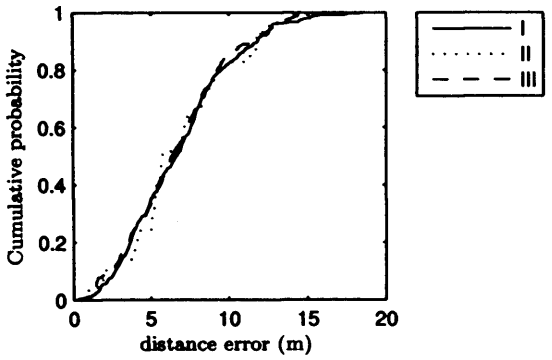
an average transmitter range of 18 m (located at the centre of the circular path). The lower location accuracy of topology four is caused by the spatial arrangement of the three readers deployed in the adjacent corners of the room in relation to the movement and pointing angle of the transmitter. This can be explained by referring once again to Figure 5.29(d), for half of the points on the circular path taken by the transmitter the line of sight signal is only available to a maximum of two out of the three readers deployed in the adjacent corners of the room. For the other half of points along the circular path the line of sight signal is only present at one of the readers. Topology one that contains three readers, each with four receivers, mounted centrally also suffers from the absence of line of sight. This topology which is illustrated in Figure 5.29(a) has the worst accuracy of the four topologies investigated with an absolute mean location error of 1.77 m and a standard deviation of 7.10 m for an average transmitter range of 18 m. Over all points along the circular path taken by the transmitter only one of the readers has a line of sight to the transmitter. The location errors are exacerbated further for this topology because the direct line of sight signals are received at large elevation angles due to the movement of the transmitter being very close to the readers in this topology. The much larger location errors encountered with topology one and four are caused by the angular samples bearings from each reader being heaviliy biased away from the true bearing. Even though the averaged angular bearings from each individual reader contain angular errors that are not multiplicative, the application of averaging algorithms II and III cannot improve the mean location error and the standard deviaton a great deal beyond that of algorithm I, particularly with topology one.

The aim of this section has been to establish the performance of the reader topologies under the most difficult propagation conditions imaginable by simulating the movement of the transmitter along a circular path whilst at the same time varying the pointing angle to be in line with the movement i.e. tangential to the current position on the circular path. From the results presented it seems that under these conditions the location error after averaging is most reduced by employing a reader topology, such as used in topology two, with a large number of equispaced readers in

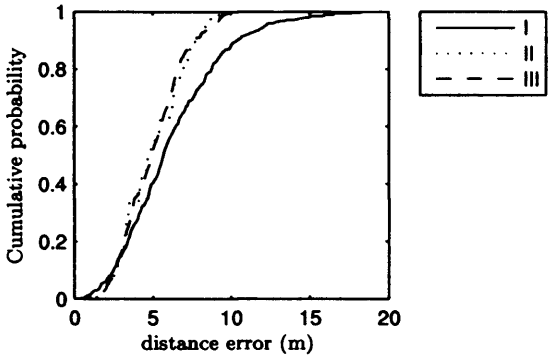
a cellular arrangement. Significantly the mean location errors from topology two are close to zero therefore the location estimates, obtained from the previous smoothing operation, could be smoothed further to improve the standard deviation of the location error. The averaging algorithms II and III applied here are more effective when (1) the reader topology contains readers with sufficient separation from the transmitter location to account for the elevation deviation between the transmitter mounting and reader mounting *and* (2) the line of sight signal is received by at least two readers over all points on the path taken by the transmitter.

Table 5.12: Mean and standard deviation of the location error (m) – Absolute , Δ_x and Δ_y – for reader topologies using model B with a reflective wall surface defined as $\varepsilon_r = 4, \sigma = 1e-8$

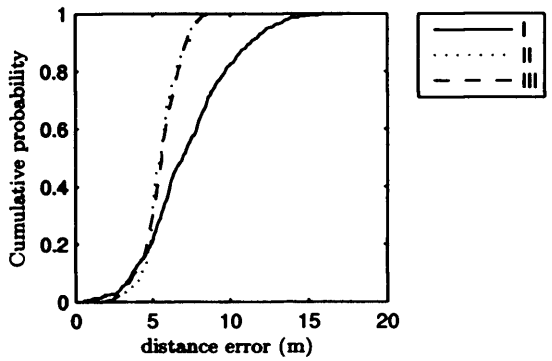
Algorithm	absolute error		error in x Δ_x		error in y Δ_y		valid samples
	μ	σ	μ	σ	μ	σ	
Topology 1							
I	1.87	7.40	1.18	5.10	-1.45	5.37	329/377
II - $\eta = 10$	2.23	7.30	1.87	4.98	-1.23	5.35	29/38
III - $\eta = 10$	1.77	7.10	1.23	4.73	-1.27	5.29	280/367
II - $\eta = 5$	2.25	7.55	1.95	5.06	-1.11	5.60	68/76
III - $\eta = 5$	2.08	7.56	1.49	5.38	-1.45	5.31	316/372
Topology 2							
I	0.28	6.96	-0.0045	5.40	0.28	4.40	376/377
II - $\eta = 10$	0.31	7.01	0.28	5.29	0.13	4.60	38/38
III - $\eta = 10$	0.23	5.34	0.0062	4.00	0.23	3.53	367/367
II - $\eta = 5$	0.28	7.05	0.53	5.29	0.32	4.66	76/76
III - $\eta = 5$	0.29	5.70	-0.012	4.28	0.29	3.76	372/ 372
Topology 3							
I	2.44	7.66	1.38	5.46	-0.39	5.38	373 /377
II - $\eta = 10$	1.92	7.73	1.86	5.66	0.49	5.25	38 / 38
III - $\eta = 10$	1.17	5.64	1.13	4.05	-0.29	3.92	367/ 367
II - $\eta = 5$	2.02	7.50	1.77	5.59	0.99	5.01	76 / 76
III - $\eta = 5$	1.18	5.87	1.14	4.24	-0.30	4.07	372 / 372
Topology 4							
I	1.54	8.56	-1.33	6.02	0.79	6.08	341/377
II - $\eta = 10$	2.31	8.13	-2.25	5.99	0.51	5.50	37 / 38
III - $\eta = 10$	1.44	6.26	-0.68	4.13	1.27	4.71	365 / 367
II - $\eta = 5$	2.54	7.91	-2.21	5.96	1.26	5.20	75/ 76
III - $\eta = 5$	1.44	6.70	-0.79	4.58	1.20	4.89	367/372



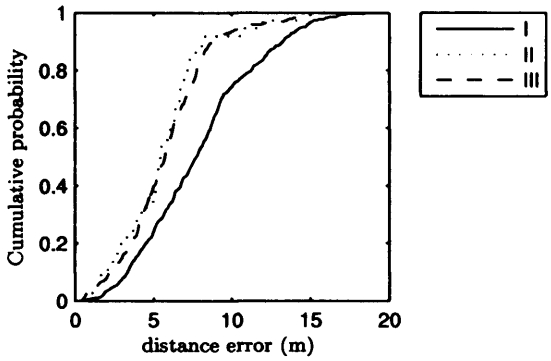
(a) Topology one - algorithms I, II and III



(b) Topology two - algorithms I, II and III



(c) Topology three - algorithms I, II and III



(d) Topology four - algorithms I, II and III

Figure 5.30: Comparison of the cdfs of the absolute location error for topologies one, two, three and four using model B with an algorithm block size, η, of 10 and a reflective wall surface defined as $\varepsilon_r = 4, \sigma = 1e-8$

5.6 Potential for multipath resolution and mitigation

Super-resolution could be considered as a possible method for separation of the direction of arrival of multi-path wave components on a receiving array. In addressing this we must first note that the multi-path components can be considered to be perfectly coherent because perfect coherence represents a difficult case for many resolution algorithms, in that the advantages of complex fading can not be exploited. The resolution performance of super resolution estimators such as MUSIC are poor particularly at low SNR because the rank of the source signal covariance is reduced by the affect of the coherent signals which impinge on the linear uniform array [36]. This means that if there were two coherent rays impinging on an array they may not be resolved, particularly if they are closely spaced in arrival angle. The spectral estimator is likely to produce a peak at an intermediate angle, approximately a weighted mean determined by the amplitudes of the two rays. One method of overcoming the singularity in the covariance matrix is spatial smoothing [40]. This is based on averaging the covariance matrix of identical overlapping arrays and requires an array of identical elements built with some form of periodic structure, such as a uniform linear/circular array. In this thesis the direction finding apparatus forms only a small part of an integrated receiver for identification and detection of a moving transmitter in real time. The use of super-resolution techniques would be inappropriate and add unnecessary complexity since a minimum of eight elements provides sufficient resolution to only identify two coherent paths. Even though we cannot resolve multipath components with a four element array, it is useful to know under what circumstances the largest azimuthal angular error occurs whenever a direct ray and a single multipath ray are received by adjacent antenna elements located on the passive direction finding array.

The received voltage on each antenna element on the passive direction finding array can be given as

$$V_r = \sum_{i=1}^{2} \frac{\left|\vec{E}_i\right|}{AF\left(\phi_i\right)} e^{j\varphi_i} \tag{5.49}$$

where: $i = 1,2$ are the direct and multipath rays respectively, ϕ_i is the angle of arrival measured relative to the boresight on antenna element one, $\vec{E}_i$ is the received electric field vector, $AF(\phi_i)$ is the antenna factor, and $e^{j\varphi_i}$ is the electrical phase of the received signal. The received voltage can also be defined using $V(\phi_i)$ which is the normalised voltage gain (field pattern) and is now given as

$$V_r = \frac{9.73}{\lambda} \sum_i \left|\vec{E}_i\right| V(\phi_i)\, e^{j\varphi_i} \tag{5.50}$$

The normalised azimuthal field pattern of the antenna element on receiver channel one has a $\cos\phi$ relationship and on the antenna element on receiver channel two it has a $\sin\phi$ relationship. The worst case angular error between adjacent antenna elements can be examined using a two ray model which contains a direct component and a multipath component. Hence the received voltages for the adjacent antenna elements squinted at 90° can be expressed as

$$\begin{aligned} V_{c1} &= \frac{9.73}{\lambda}\left[A_d \cos\phi_d e^{j\varphi_d} + A_m \cos\phi_m e^{j\varphi_m}\right] \\ V_{c2} &= \frac{9.73}{\lambda}\left[A_d \sin\phi_d e^{j\varphi_d} + A_m \sin\phi_m e^{j\varphi_m}\right] \end{aligned} \tag{5.51}$$

where: A_d, ϕ_d, φ_d are the electric field strength, arrival angle and electrical phase of the direct ray and A_m, ϕ_m, φ_m are the electric field strength, arrival angle and electrical phase of the multipath ray. The maximum error in the received antenna amplitudes, V_{c1} and V_{c2}, occurs when the relative electrical phase between the direct ray and the multipath ray are either in-phase, $\varphi_d - \varphi_m = 0$, or out of phase $\varphi_d - \varphi_m = \pm\pi$. Both conditions are equally valid, here we take the in-phase condition and the antenna element voltages simplify to

$$\begin{aligned} V_{c1} &= \frac{9.73}{\lambda}\left[A_d \cos\phi_d + A_m \cos\phi_m\right] \\ V_{c2} &= \frac{9.73}{\lambda}\left[A_m \sin\phi_d + A_m \sin\phi_m\right] \end{aligned} \tag{5.52}$$

The antenna element voltages are amplified by the respective receiver channels and a signal strength is then estimated using the control voltage on the AGC. As derived in Chapter 4 the logarithmic voltage ratio between the field patterns of the adjacent channels is a $\tan\phi$ function. Therefore the angular error in the angle of arrival caused by the multipath can be calculated in terms of the voltage ratio V_{c1}/V_{c2}. Using this ratio the angular error in the measurement of the angle of arrival for balanced channels in the absence of additive Gaussian noise can be defined as

$$\begin{aligned}
\varepsilon &= \tan^{-1}(V_{ratio}) - \phi_d \\
&= \tan^{-1}\left(\frac{A_d \sin\phi_d + A_m \sin\phi_m}{A_d \cos\phi_d + A_m \cos\phi_m}\right) - \phi_d \\
&= \tan^{-1}\left(\frac{\sin\phi_d + \frac{A_m}{A_d}\sin\phi_m}{\cos\phi_d + \frac{A_m}{A_d}\cos\phi_m}\right) - \phi_d
\end{aligned} \tag{5.53}$$

Using Equation 5.53, the angular error is calculated whenever the three parameters A_m/A_d, ϕ_d and ϕ_m are varied. Here the angular error is evaluted for three receive angles of 0°, 45°, and 90°, within the quadrant common to both elements and for the three multipath ray amplitudes of 0 dB, -3 dB and -20 dB relative to the amplitude of the direct ray. These relative amplitudes correspond to the additonal distance travelled by the multipath rays of 0 m, $\sqrt{2}$ m and 10 m since the attenuation in the electric field is proportional to the inverse of this distance. The results are shown in Table 5.13. Unsurprisingly a zero angular error occurs when the angle of arrival of the multipath is conincident with the angle of arrival of the direct ray. The largest angular error of $\pm$ 22.5° occurs when the angle of arrival of the multipath ray is orthogonal to the angle of arrival of the direct ray when have equal amplitudes i.e. in this case when the rays are impinging at adjacent antenna element boresights.

	$\frac{A_m}{A_d}$	$\phi_m = 0°$	$\phi_m = 45°$	$\phi_m = 90°$
$\phi_d = 90°$	1(0dB)	-22.50	-15	0
	$\frac{1}{\sqrt{2}}$(-3dB)	-20.03	-13.53	0
	$\frac{1}{10}$(-20dB)	-8.77	-6.11	0
$\phi_d = 45°$	A(0dB)	0	22.5	7.5
	$\frac{1}{\sqrt{2}}$(-3dB)	6.11	11.25	6.11
	$\frac{1}{10}$(-20dB)	1.59	1.43	1.3
$\phi_d = 0°$	1(0dB)	undefined	22.5	22.5
	$\frac{1}{\sqrt{2}}$(-3dB)	0	18.24	20.03
	$\frac{1}{10}$(-20dB)	0	6.46	8.77

Table 5.13: Maximum angular error for a two multipath model applied to a pair of antenna elements squinted by 90° in azimuth

5.6.1 Multipath compensation

A possibility for removing multipath error from the angular bearing errors at the readers is to the measure or model the error statistics of the transmitter as it moves in fixed steps on a grid within the room. Each position in the room is represented by a bias and an associated variance value which is stored in a lookup table. Alternately, the parameters could also be represented as multipath function by means of a series of polynomials and/or harmonic terms. This approach is attractive because the specification of the approximation function requires less memory than the specification of a look-up table. If the multipath shows variability, then it may be preferable to have a grid of non-uniform mesh size. Although this would make the search slower and consume more memory. The cell size within the mesh depends on the velocity of the variation of multipath errors with respect to the location of the target. Small cells give high accuracy but also require more memory to store the look-up table. At each cell, the multipath could be compensated by removing the bias from the current average bearings using the stored bias and variance for the current estimated location. The method would work if the variables of the measurements do not change and could conceivably perform very well when there is no height difference between the transmitter and readers and the tranmitter contains

an omni-directional beam pattern.

5.7 Implications for the reader network performance

Unless the tags have antennas both on the chest and the back of the person, the number of readers may have to be nearly doubled to cater for the tag directivity. Although the angular error is reduced by ensuring a sufficient number of readers receive a LOS signal this does not prevent angular errors caused by multipath propagation. First of all multipath fading could cause a tag to be temporarily in a blind spot. However if the three or four distinct reader are arranged to keep all relevant areas within the field of view, total loss of cover (by a combination of fading, obstruction and unfavourable location) will be rare. Furthermore, with a wavelength of 6cm, even a stationary tag wearer will rarely keep the tag within a fade for more than some seconds. A more serious effect of multipath propagation is that it can result in false bearings. In particular, if the direct signal is faded or obstructed, its place may be taken by quite a strong wall-reflected signal. With just two bearings, this will produce a false location. With three bearings, it will produce one true and two false locations, with little direct indication which is which. Although the pair of readers containing receivers with the strongest signals can probably be regarded as being more likely to represent true direct signals. The overall implication of all this is that 'individual' locations are fallible, and tracking must be an integrated process, taking into account continuity of location, limitations of velocity, physical constraints of movement on the ground, distinct confidence ratings of individual observations/locations and so forth.

6. EXPERIMENTAL RESULTS - FIELD TRIALS

This chapter presents experimental results taken in the departure lounge in Debrecen airport, Hungary and at a second location within UCL.

6.1 Site A - Debrecen airport trials

The experiment involved recording the signal strengths on the four channels of a reader received from a transmitter which was moved in pre-determined positions and orientations. The measurements for each test were then post processed to quantify the angular error performance of the reader. The angular error could then be extrapolated to provide an indication of location error assuming a distance times tan(angle)law. The tests chosen were based on ideas developed earlier in the thesis. First of all, it was imperative to to test the dependence of angular performance of the reader on its position. This would provide evidence supporting the hypothesis that a corner-mounted base station can provide a lower RMS angular deviation from the mean error than a base station situated in an open central location because there are no quadrant errors apparent in the corner mounted base station. To obtain a meaningful performance comparison, the signal strength measurements were recorded in different locations, firstly in a open location using four channels where the AOA was computated using the strongest adjacent channel algorithm, and secondly in a corner location using only two channels with the AOA computed directly. In each reader location, the mean and rms variation of the AOA bearing errors were computed for a matrix of transmitter positions with each location having four pointing orientations. As well as those tests, it was also necessary to examine the sensitivity of a reader channel in the open location. This capability was performed in a number of ways that included RSS/range measurements, by also placing the tag on the ground antenna face down near the reader and noting the received signal, by taking the tag into other rooms which are at far distances from the base station in the departure area, for instance a toilet, and noting the received signal. The preceeding tests can not solely describe the expected performance expected in

the indoor environment without additonal tests which quantify the effect of people on the propagation characteristics in the room. There are two equally valid tests which can provide useful results on this matter. First, the effect of obstruction on the received signal strength on one channel was implemented by positioning a number of people, equispaced, in a line between the base station and the transmitter. Second, the effect of crowding was implemented by positioning a number of people in a square around the transmitter position. The received signal strengths on all channels and the subsequent calculated angle could then be compared with the previous measurements assuming the transmitter was positioned and orientated in an equivalent direction. Both scenarios mimmic the crowding scenario ever-present at many large airports. In short, the tests provide important experimental results regarding the performance of an active RF-ID system operating at 5.8 GHz in a modern airport departure area. The tests are also affected to a large extent by the topology of the indoor environment and this must be considered so that most appropriate locations for the corner and centrally located are found. The topology is therefore considered in detail in the section to follow.

6.1.1 Description of departure area

The type of room topology can have a significant impact on the location performance of the system. Above all else, the quality of the bearing estimates from each reader is affected by a lack of a direct line of sight signal but this is exacaberated further by reflectors within the room. These reflectors may be either non moveable features perhaps shutters, walls and/or concrete support columns and other items which consist only partially of metal reflective surface or other moveable items which support reflections, for instance passenger seating. Accordingly, the range and accuracy of the location system is dependent on the number of readers and their individual positions. The placement of a base station within the location system must be carefully considered to limit the disturbance to the bearing estimate caused by potential reflectors. However if the reflectors are too numerous then placement at a reasonable distance from all reflectors is not always possible. A reader placed in the corner obviates the effect of the surrounding walls, but has limited range

and angular bearing accuracy if there are intervening structures such as columns. Elsewhere, a reader placed nearer the middle of the room has wider coverage but the bearing accuracy is severely affected by the walls and other artifacts; although the accuracy could be improved by adjusting the number of adjacent channels used to compute the bearing for example if it is known that the artifact is directly in line with one of the receive channels. As consequence of removing a channel from the bearing calculation, as would happen if a base station is placed near an artifact like a building support column, a dead zone would be present in that 180 degree sector. Since the base station can be placed in any number of positions, the deployment of the base stations is aided by using a survey of the room to identify the critical features of the room topology which may affect the performance of the system.

The plan view of the departure area can be seen in Figure 6.2. The working area is defined within the region marked by a blue line in which the maximum room length is 20 m and the maximum room width is 13.8 m. The working area has four main sides, the left side of the room faces onto the airfield. It is composed of panes of glass surrounded by rectangular metallic banding. The top side is made of brick whereas the right side contains three identical metal shutters separated by small sections of brick work. The storage at the bottom right hand corner of the room is surrounded by brick wall and the remaining side at the bottom is made of wood which encloses a smoking waiting room. It is located near the main brick wall enclosing the room. The floor which consists of square tiles of size 40.1 cm and contains a large number of moveable seats positioned on the departure side of the room which can be seen in Figure 6.1. The ceiling which can also be seen is 2.9 m in height and consists of an array of drop-in polystyrene ceiling tiles. The rest of the floor area is open space mainly near the entrances to toilets at the top of the room and the right hand side of the room ; there are three pillars of square size 45 cm which are positioned on a axis running centrally down the room and are separated from each other by 5.74 m.

Figure 6.1: The departure area with moveable seating and view onto airfield

6.1.2 Description of experiments

The reader used in the experiments was located in the positions denoted by a filled black square in the floor plan shown in Figure 6.2. These were chosen in order to show comparative experimental results between two different deployments. A series of measurements were conducted and these are briefly discussed now. Firstly, the location accuracy was measured using all four sectors of the tag reader. The tag reader was placed in a central portion of the departure lounge and a member of the Optag team wore a tag and stood at pre-determined locations on a grid. The furniture in the room was moved to allow the measurement of a regular grid with origin located at the base station. This consisted of a total of 54 measurement positions on a 9 by 6 grid. At each position, the reader signal information was

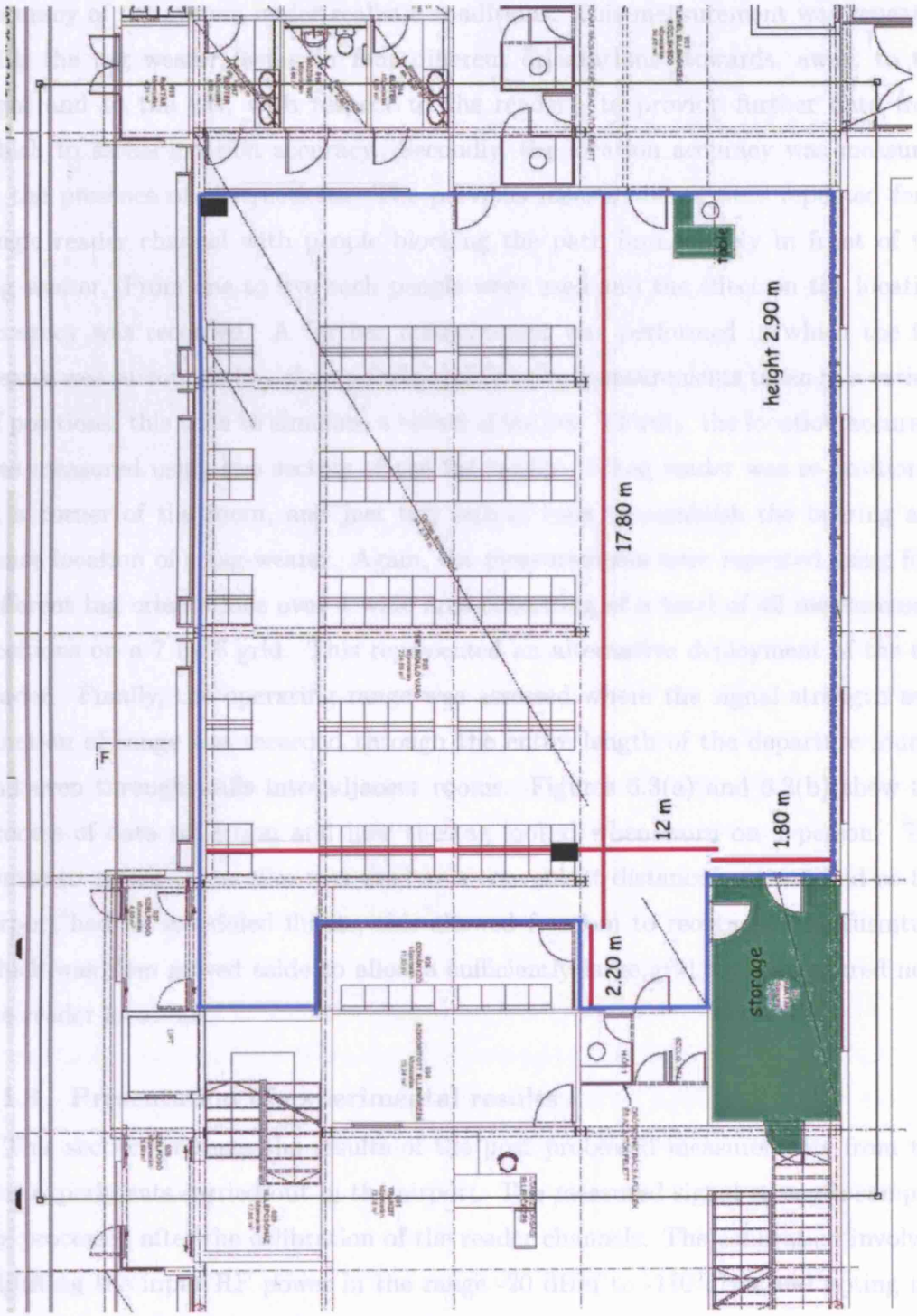


Figure 6.2: A plan view of the departure lounge

recorded and subsequently correlated with the actual position, in order to assess the accuracy of the system under realistic conditions. This measurement was repeated with the tag wearer facing in four different orientations (towards, away, to the right and to the left, with respect to the reader), to provide further data from which to assess location accuracy. Secondly, the location accuracy was measured in the presence of obstructions. The previous measurements were repeated for a single reader channel with people blocking the path immediately in front of the tag wearer. From one to five such people were used and the effect on the location accuracy was recorded. A further measurement was performed in which the tag wearer was surrounded by eight people and location measurements taken in a variety of positions, this time to simulate a crowd situation. Thirdly, the location accuracy was measured using two sectors of the tag reader. A tag reader was re-positioned in a corner of the room, and just two sectors used to establish the bearing and hence location of a tag-wearer. Again, the measurements were repeated using four different tag orientations over a wide area consisting of a total of 42 measurement positions on a 7 by 6 grid. This represented an alternative deployment of the tag reader. Finally, the operating range was assessed where the signal strength as a function of range was recorded through the entire length of the departure lounge and even through walls into adjacent rooms. Figures 6.3(a) and 6.3(b) show the process of data collection and how the tag looked when worn on a person. The centre to centre of the tiles was used as a convenient distance measure and as the airport had no scheduled flights, this allowed freedom to reorganise the furniture which was then moved aside to allow a sufficiently large grid to be measured near the reader locations.

6.1.3 Presentation of experimental results

This section presents the results of the post processed measurements from the four experiments carried out in the airport. The measured signal strength samples are processed after the calibration of the reader channels. The calibration involved adjusting the input RF power in the range -20 dBm to -110 dBm and noting the changes in the AGC control voltage. The calibration identified any voltage offsets

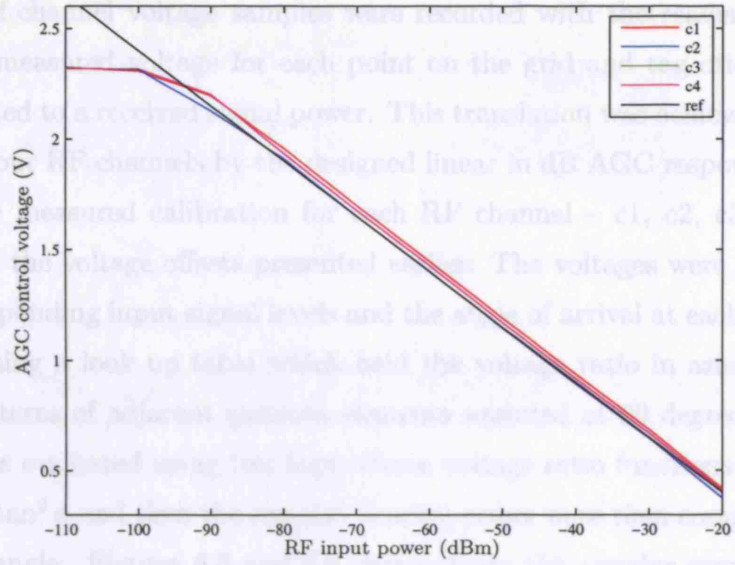
(a) measurement grid for central-mounted reader

(b) tag worn using lanyard

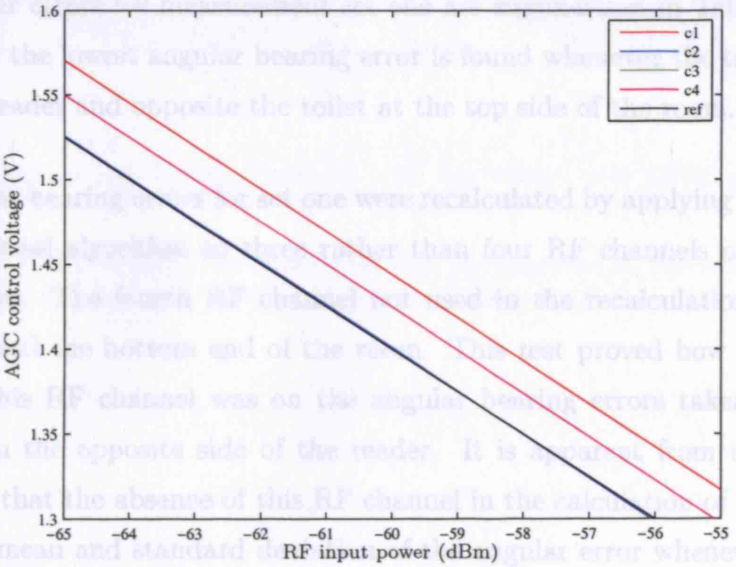
Figure 6.3: A typical measurement scenario

from the designed linear response of 40 dB per volt slope and a calibrated reference voltage of 0.65 volts for -30 dBm input signal power. The calibration also ensured that any voltage offsets present between channels are accounted for whenever the AOA is computed from the voltage difference between the two strongest adjacent channels. If there are unknown voltage offsets present, these add error to the calculation of the angular bearing since there is now channel unbalance. This potential angular error can be quantified by an understanding of the variation of the azimuthal voltage ratio versus angle of arrival. First of all, if the logarithmic $\tan\phi$ voltage ratio is linearised in the range 15° to 75°, the function can be approximated by a slope of gradient 0.37 dB per degree as shown in Chapter 4 Equation 4.23. This slope shows the sensitivity of the angular bearing to a difference in the voltage ratio between the adjacent RF channels. Hence a voltage offset applied to an RF channel of say 10 mV relates directly to a signal ratio offset of 0.4 dB and this in turn leads to an 1° angular error when all other things are taken equal. As a consequence the calibration must be considered as the first step in the accurate post processing of the measured data otherwise a systematic error is automatically incurred.

The results of the calibration are shown in Figure 6.4. The calibration indicates that two of the channels closely matches the desired characteristic. The others are offset by 0.043 and 0.024 volts respectively (corresponding to 1.72 and 0.96 dB).



(a) whole range



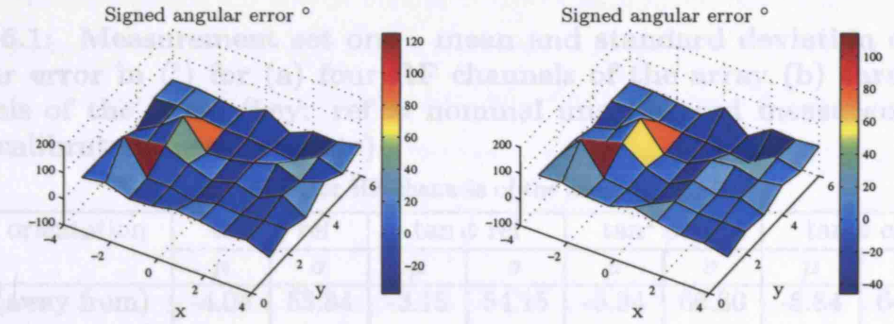
(b) central portion

Figure 6.4: Plots of the channel calibration for the four receiver channels in the reader used in the experiments in site A with respect to the reference response with the (signal strength(nominal) = −(AGCControlVoltage − 0.65) ∗ 40 − 30 dBm)

6.1.4 Measurement set one - central reader location

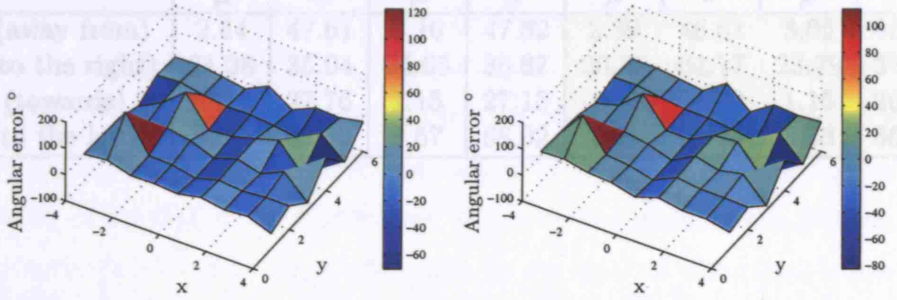
A series of channel voltage samples were recorded with the reader located centrally. The measured voltage for each point on the grid and tag orientation were then translated to a received signal power. This translation was achieved by approximating all four RF channels by the designed linear in dB AGC response or exactly by using the measured calibration for each RF channel – c1, c2, c3, c4 – which incorporated the voltage offsets presented earlier. The voltages were then changed to the corresponding input signal levels and the angle of arrival at each position was computed using a look up table which held the voltage ratio in azimuth between the field patterns of adjacent antenna elements squinted at 90 degrees. The angle of arrival was evaluated using two logarithmic voltage ratio functions $20\log_{10}\tan\phi$ and $20\log_{10}\tan^2\phi$ and then the angular bearing errors were then computed relative to the true angle. Figures 6.5 and 6.6 demonstrate the angular error for the two functions with and without calibrated channels. The mean and standard deviation of the angular errors for measurement set one are summarised in Table 6.1(a) and indicate that the lowest angular bearing error is found whenever the tag is pointing toward the reader and opposite the toilet at the top side of the room.

The angular bearing errors for set one were recalculated by applying the strongest adjacent channel algorithm to three rather than four RF channels on the central located reader. The fourth RF channel not used in the recalculation was near a pillar and faced the bottom end of the room. This test proved how sigificant the absence of this RF channel was on the angular bearing errors taken from a tag positioned on the opposite side of the reader. It is apparent from the results in Table 6.1(b) that the absence of this RF channel in the calculation of the AOA has reduced the mean and standard deviation of the angular error whenever there is a lack of a direct line of sight to the reader. The lowering in error can be seen more clearly in the highlighted results using the calibrated RF channels. As expected the angular accuracy is very similar for both applications of the strongest adjacent receiver algorithm when the tag is pointing towards the reader because a line of sight signal is available.



(a) log voltage ratio $(20 \log_{10} \tan \phi)$ ref (b) log voltage ratio $(20 \log_{10} \tan^2 \phi)$ ref

Figure 6.5: The signed angular error for measurement set one using four uncalibrated RF receivers when the tag is pointing towards the reader



(a) log voltage ratio $(20 \log_{10} \tan \phi)$ cal (b) log voltage ratio $(20 \log_{10} \tan^2 \phi)$ cal

Figure 6.6: The signed bearing error for measurement set one using four calibrated receivers when the tag is pointing towards the reader

Table 6.1: Measurement set one - mean and standard deviation of the angular error in (°) for (a) four RF channels of the array (b) three RF channels of the array (key: ref = nominal uncalibrated measurements, cal = calibrated measurements)

(a) four RF channels of the array

Tag orientation	$\tan^2\phi$ ref		$\tan\phi$ ref		$\tan^2\phi$ cal		$\tan\phi$ cal	
	μ	σ	μ	σ	μ	σ	μ	σ
+y (away from)	-4.04	53.84	-3.15	54.15	-6.34	66.80	-5.84	64.87
+x (to the right)	21.21	43.57	24.50	47.26	22.64	48.27	24.68	48.29
-y (towards)	-0.14	27.76	1.15	27.17	-0.10	27.22	1.15	26.64
-x (to the left)	0.09	66.30	7.91	70.82	-5.15	69.62	9.10	72.87

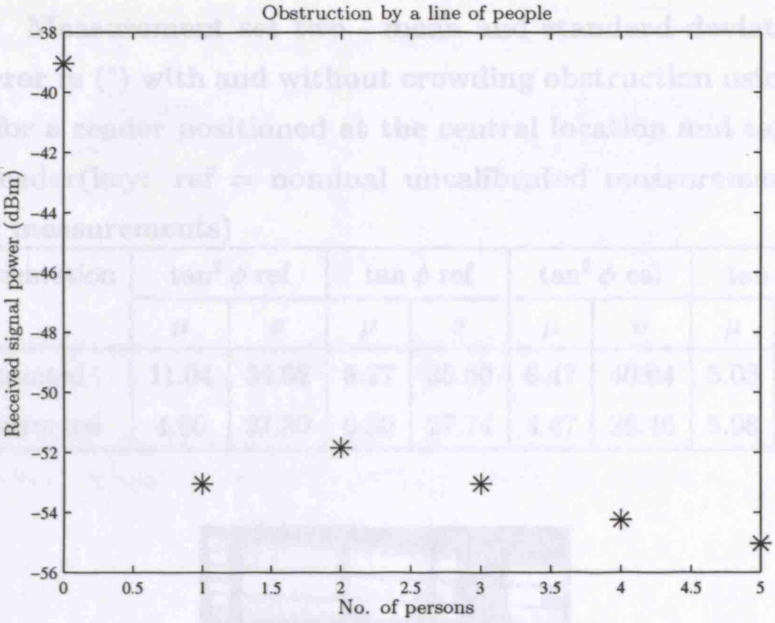
(b) three RF channels of the array

Tag orientation	$\tan^2\phi$ ref		$\tan\phi$ ref		$\tan^2\phi$ cal.		$\tan\phi$ cal.	
	μ	σ	μ	σ	μ	σ	μ	σ
+y (away from)	2.44	47.61	3.16	47.82	2.39	46.61	3.06	46.03
+x (to the right)	21.38	35.04	24.65	38.87	20.85	34.17	23.79	37.27
-y (towards)	-0.14	27.76	1.15	27.15	-0.10	27.22	1.15	26.62
-x (to the left)	7.87	62.32	8.67	68.02	7.66	61.29	8.33	66.21

6.1.5 Measurement set two - signal obstruction at central reader location

A series of channel voltage samples were recorded on a single RF channel with the direct path from the transmitter to the reader obstructed by people in line. The voltage were translated to signal strengths in the method described in measurement set one. The first test measured the received signal strength on the boresight of one channel in the reader when obstructed by five people in line. Each person was separated by two tiles in a line directly in front of the tag. After each measurement the person at the front of the line nearest the reader was removed and another measurement recorded. The results shown in Figure 6.7(a) indicate that the signal strength is affected primarily by the first person blocking the transmitted signal with subsequent people causing a mere +3 dB of additional attenuation. A second test was conducted in which again a series of channel voltage samples were taken but now using four RF channels in the reader with the direct path and multipath from the transmitter to the reader obstructed by people in a crowd around the transmitter. This test utilised the same setup as in measurement set one. The voltages were translated to signal strengths in the method described before. The affect of crowding was then tested using eight people surrounding the tag. The results were taken on a much reduced sample set than used in set one. It consisted of a total of 15 measurement positions on a 5 by 3 grid with now a larger step of four tiles taken within the original grid ie range x = -4:2:4 and y = 2:2:6. This was then compared directly with the corresponding unobstructed measurements from earlier measurements at these positions. Figure 6.8 shows the positioning of the people surrounding the tag wearer in a rectangular arrangement with each person separated by at least one and half tile lengths or 60.075 cm. The angular errors from the crowding obstruction measurements are compared with the angular errors from the unobstructed measurements as shown in Table 6.2. The angular errors for both scenarios are also plotted for a $20 \log_{10} \tan^2 \phi$ voltage ratio as shown in in Figures 6.9. The results from each function indicate that the angular error for the unobstructed samples is lower than the obstructed samples at the same measurements points both for the calibrated and uncalibrated channel approximation. If we average

the errors from each function we find that the crowding measurements have an angular mean error, μ, of 7.70° and a standard deviation, σ, of 37.03° whereas the unobstructed measurements have lower error with an angular mean error, μ, of 5.33° and a standard deviation, σ, of 27.01°.



(a) Variation of the received signal strength with the number of persons obstructing the LOS

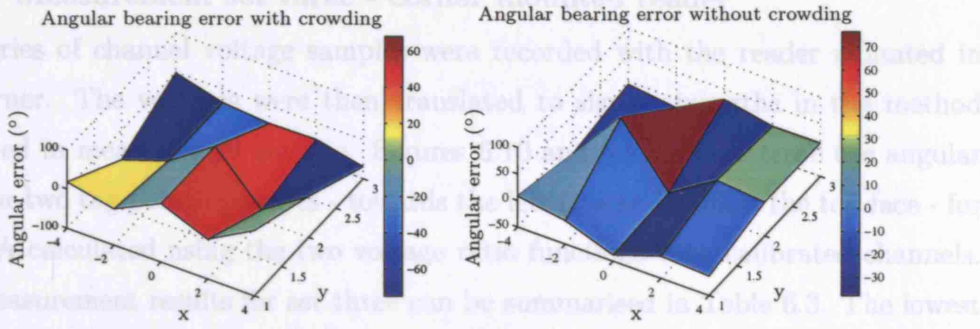
(b) Line obstruction of five persons in front of the reader

Figure 6.7: Effect of a line of people on received signal strength at a single receiver channel in the reader

Table 6.2: Measurement set two - mean and standard deviation of the angular error in (°) with and without crowding obstruction using four RF receivers for a reader positioned at the central location and tag pointing towards reader(key: ref = nominal uncalibrated measurements, cal = calibrated measurements)

Tag orientation	$\tan^2\phi$ ref		$\tan\phi$ ref		$\tan^2\phi$ cal		$\tan\phi$ cal	
	μ	σ	μ	σ	μ	σ	μ	σ
obstructed	11.04	34.92	8.27	35.50	6.47	40.04	5.03	38.76
unobstructed	4.60	27.30	6.30	27.74	4.47	26.46	5.98	26.54

Figure 6.8: Some of the crowd of 8 surrounding the tag wearer hidden from view in the middle



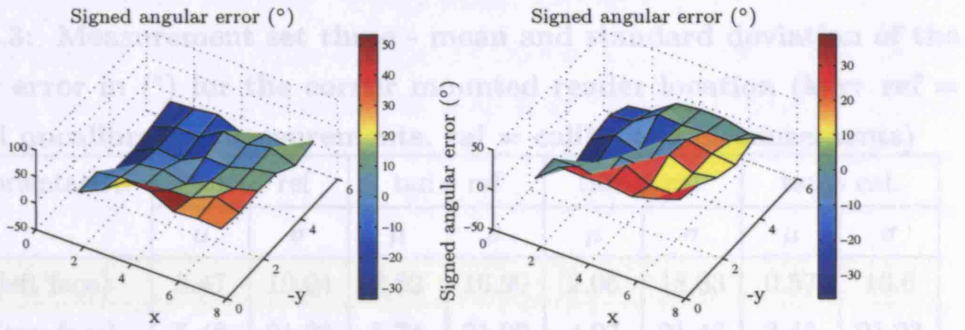
(a) with obstruction - log voltage ratio ($20\log_{10}\tan\phi$) cal

(b) without obstruction - log voltage ratio ($20\log_{10}\tan^2\phi$) cal

Figure 6.9: The signed bearing error for the crowding test using four calibrated receivers when the tag is pointing towards the reader

6.1.6 Measurement set three - corner mounted reader

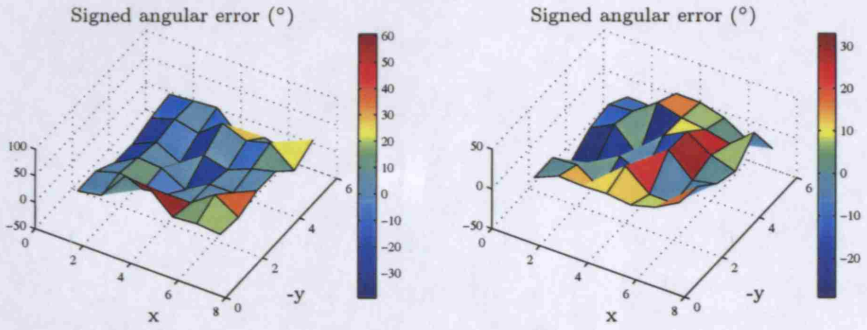
A series of channel voltage samples were recorded with the reader mounted in the corner. The voltages were then translated to signal strengths in the method described in measurement set one. Figures 6.10 and 6.11 demonstrate the angular error for two tag pointing angles - towards the left face and toward the top face - for an AOA calculated using the two voltage ratio functions with calibrated channels. The measurement results for set three can be summarised in Table 6.3. The lowest angular error occurs when the tag is pointing towards the left hand side of the room as highlighted in the table. At this pointing angle the angular errors computed from the calibrated RF receiver channels also show lower bias and variance than the angular errors computed from uncalibrated RF channels. It also clear from the angular bearings errors calculated at this pointing angle that the logarithmic voltage ratio ($20\log_{10}\tan\phi$) performs better than the logarithmic voltage ratio ($20\log_{10}\tan^2\phi$). This implies that the field patterns on the antenna elements on the direction finding array fit a $\cos\phi$ pattern more closely than a $\cos^2\phi$. If we now turn our attention to the relative performance of the two reader mountings we find that the angular accuracy of the corner-mounted reader is better than the angular accuracy of the central-mounted reader for all four pointing angles. The lowest angular error for the corner-mounted reader is a mean, μ, of 0.57° and standard deviation, σ, of 16.6° and this occurs for a $\tan\phi$ cal. function when the tag is pointing toward the left face. In contrast to this result the lowest angular error for the central-mounted reader as shown in Table 6.1(a) is a mean of $-0.14°$ and a standard deviation of 27.8° and this occurs for a $\tan^2\phi$ ref. function when the tag is pointing toward the bottom face. The difference in angular performance can be seen more distinctly when the pointing angles are not favourable. The largest angular error for the corner-mounted reader is a mean of 18.93° and a standard deviation of 29.90° and this occurs for a $\tan\phi$ ref. function when the tag is pointing toward the bottom face. This is much lower than the largest angular error encountered with the central mounted reader. Again referring to Table 6.1(a) the largest angular error for the central mounted reader is a mean of $-30.71°$ and standard deviation of 68.08° and this occurs for a $\tan^2\phi$ cal. function when the tag is pointing toward the top face.



(a) tag pointing towards to the left side of the room

(b) tag pointing towards the top side of the room

Figure 6.10: The signed bearing error for measurement set three calculated from a log voltage ratio ($20 \log_{10} \tan^2 \phi$) using only two calibrated receivers in the reader



(a) tag pointing towards the left side of the room

(b) tag pointing towards the top face

Figure 6.11: The signed bearing error for measurement set three calculated from a log voltage ratio ($20 \log_{10} \tan \phi$) using only two calibrated receivers in the reader

Table 6.3: Measurement set three - mean and standard deviation of the angular error in (°) for the corner mounted reader location (key: ref = nominal uncalibrated measurements, cal = calibrated measurements)

Tag orientation	$\tan^2\phi$ ref		$\tan\phi$ ref		$\tan^2\phi$ cal.		$\tan\phi$ cal.	
	μ	σ	μ	σ	μ	σ	μ	σ
-x (left face)	3.47	19.04	2.82	16.90	2.06	18.83	0.57	16.6
+y (top face)	5.48	21.82	5.74	21.92	4.07	21.46	3.48	21.23
+x (right face)	-1.42	26.59	-5.12	27.70	-2.84	26.63	-7.50	27.95
-y (bottom face)	13.15	26.65	18.93	29.60	11.73	25.96	16.60	28.21

Figure 6.12: The corner reader mounting used in measurement set three

6.1.7 Measurement set four - range measurements

A set of channel voltages from a single RF channel was recorded whenever the transmitter was moved in a straight line away from the reader toward the other end of the room. The reader was located in the central position and voltages were converted to signal strengths as before. Referring to Figure 6.13(a) we see that the received signal strength with distance has been measured by the receiver on the central mounted reader for the same four tag orientations as used in the previous measurements. As before, the pointing angle of the tag orientations are referred to the standard right hand co-ordinate system with x axis aligned to the right face of the room whilst the y-axis is aligned away from the reader to the top of the room. Specifically 0° means the tag is orientated away from the reader and towards the top face, 180° means the tag is pointing towards the reader and also towards the bottom face, 90° means the tag is pointing toward the right face and finally -90° means the tag is pointing toward the left face. Looking now at the data we find that the signal strength rolls off much more rapidly with distance when the transmitter is not facing toward the receiver. The amplitude varies by as much as 10 dB over a distance as short as 1.6 m whenever the tag is pointing toward the right face. However when the transmitter is facing toward the receiver the received signal strength rolls off more slowly with distance and has less amplitude variation between samples as was present with the other three tag orientations. If these results were extrapolated beyond the 15 m range tested here it is very likely that the reader could identify and perform direction finding of tags located at distances as a high as 100 m. This is because the linear in dB signal strength indicator has a sensitivity of about -90 dBm.

A log distance path loss fit for the LOS range measurement (tag orientation of 180 °) could accurately define how shadowing has affected the signal propagating at a frequency of 5.8 GHz in this environment. The log distance path model is an indoor propagation that can predict the path loss a signal encounters inside a building over distance. It takes account not only of path loss which is caused by dissipation of the power radiated by the transmitter but also shadowing which is

caused by obstacles between the transmitter and receiver that attenuate the signal through absorption, reflection, scattering and diffraction. The log distance path loss model can be formally expressed as

$$L = L_o + 10\gamma \log_{10}\left(\frac{d}{d_o}\right) + \chi_g \tag{6.1}$$

where: L = the total path loss (dB), L_o is the path loss at the reference distance, γ is the path loss distance exponent, χ_g is a log normal distribution with a standard deviation σ_g representing a fading factor - shadowing fading or slow fading.

The path loss at the reference distance here is roughly calibrated and taken at two tile lengths, 80.1 cm, directly in front of the reader, because multipath effects would be expected to be generally minimal at this distance. The first model chosen to fit the data was a path loss only model which has a path loss distance exponent γ of 2 representing a 20 dB /decade rolloff in the received signal strength. As illustrated in Figure 6.13(b) this model does not fit the measured data very well and a second model was chosen to fit the data. This model used a least squares approach to fit a straight line to the measured signal strength samples and this produced a path loss distance exponent γ of 1.82 or a 18.2 dB /decade rolloff. As a consequence of this approach the second model fits the data much better. The standard deviation σ_g between the model and the measured signal strength samples can be calculated by taking the root mean square of the deviations between the estimated signal strengths of the model and the measured signals strengths [41]. It can be expressed as

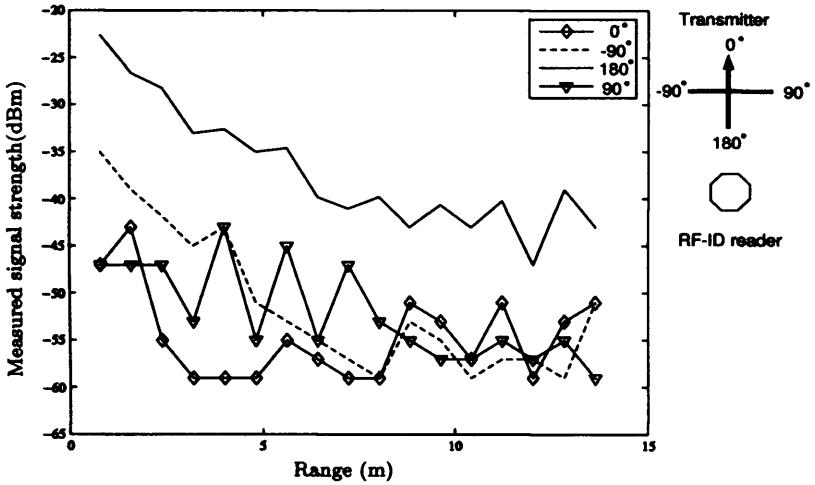
$$\sigma_g = \sqrt{\frac{1}{S}\sum_{i=1}^{S}\left[\mathrm{P_{measured}}\left(d_i\right) - \mathrm{P_{model}}\left(d_i\right)\right]} \tag{6.2}$$

where S is the total number of received signal strength samples, $\mathrm{P_{measured}(d_i)}$ is the measured received signal strength at distance i and $\mathrm{P_{model}(d_i)}$ is the modeled received signal strength. For model two this provides a σ_g of 2.0549 dB from the 17 received signal strengths. The coefficients for model two of 10γ = 18.20 dB and σ_g = 2.0549 dB can be compared with previous experimentally derived coefficients as

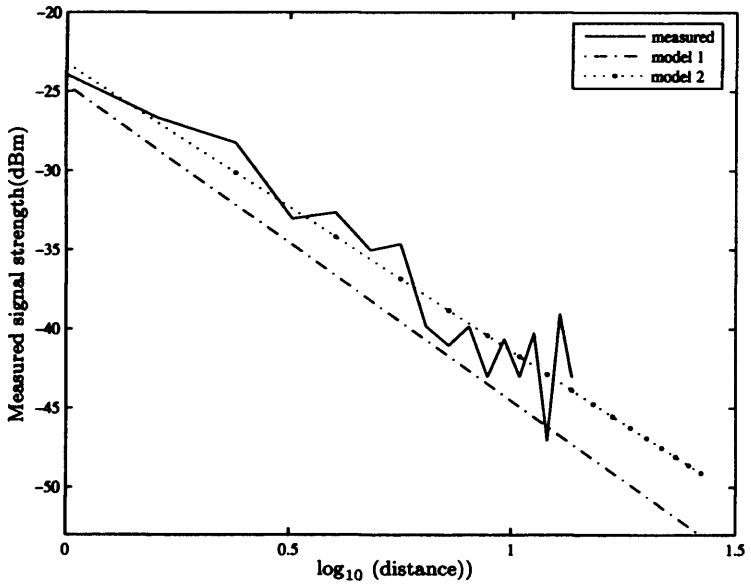
shown in Table 6.4[42].

Table 6.4: Path loss models measured for other building types

Building type	Frequency of transmission	10γ	σ_g (dB)
Retail store	914 MHz	22	9.7
Grocery store	914 MHz	18	5.2
Office with hard partition	1.5 GHz	30	7
Office with soft partition	900 MHz	24	9.6
Office with soft partition	1.9 GHz	26	14.1
Textile or chemical	1.3 GHz	20	3.0
Textile or chemical	4 GHz	21	7.0, 9.7
Metalworking	1.3 GHz	16	5.8
Metalworking	1.3 GHz	33	6.8



(a) measured received signal strength with distance



(b) log path loss models fitted to the line of sight transmitter pointing angle of 180 °

Figure 6.13: Range measurements from the transmitter to a single receiver on the central mounted reader for four tag antenna orientations

6.1.8 Measurement analysis

The section provides a summary of RF-ID tag location results obtained from trials at site A in Debrecen airport. There are several key conclusions to be drawn from the measurements. Location accuracy with a centrally-mounted reader is very variable, typically of the order of 3 - 5 metres. The results show that propagation effects are very significant in this particular reader configuration, as reflections from walls and other objects can often exceed the strength of the direct path, with misleading results. This is not at all surprising as amplitude-comparison direction finding is inherently vulnerable to these effects. The best results are obtained when the tag is approximately facing a given reader. However the location accuracy at other tag orientations can be improved slightly as seen in Table 6.1(b) by excluding the RF channel on the other side of the reader when calculating the angle of arrival using the strongest adjacent signal algorithm. Operating range was found to greatly exceed the required 10 m. The signal was well above threshold even to ranges of 25 m and with an intervening wall (between the departure lounge and toilet). The signal was only lost at such ranges when the tag was placed face-down on the floor. This is a very pleasing result, indicating that quite large cell sizes should be possible, with reduced infrastructure requirements. Obstruction by one or more persons causes a significant reduction in signal strength - around 10 dB, but still provides ample signal. It tends to reduce location-estimate accuracy, but not significantly as seen in the crowding comparison in Table 6.3. The presence of more than one person in the direct path makes little difference to the received signal as seen in Figure 6.7(a), indicating that at these frequencies (5.8 GHz) a single person causes a large attenuation of the direct path. Results obtained with the readers placed in the corners of the room were much better in terms of direction-finding and hence location accuracy. This is doubtless due to the improved propagation environment with a corner-mounted reader. In addition, this arrangement makes for a more straightforward location algorithm. This configuration is therefore preferable. One of the results obtained with corner-mounting is shown in Figure 6.11(a). This result indicates the bearing error for different positions of the tag-wearing individual, over a 6 x 7 grid (with a grid spacing of 801 mm). It is apparent that the mean error is

close to zero, indicating that there is no bias in the reader, but there is considerable variation around this mean, dependent on the tag location. The standard variation in this particular set of results is 16 degrees - equating to 1.4 m of location error at the typical operating range of 5 m.

6.2 Site B - UCL trials

This section provides additonal measurements to support the results from the Debecen trial by forming actual location errors rather than estimated location errors using experiments conducted in a UCL facility. The room used for this experiment is approximately 15 m long by 8 m wide. A representation of the major structural features of the room is shown in figure 6.14. The perimeter of the room is constructed of brick and there are two large metallic columns with brick bases situated near the centre of the room and the northern most face of the room. In addition there is a 6ft long by 8 ft high brick wall, not far from the centrally located column, that protudes from the western face of the room.

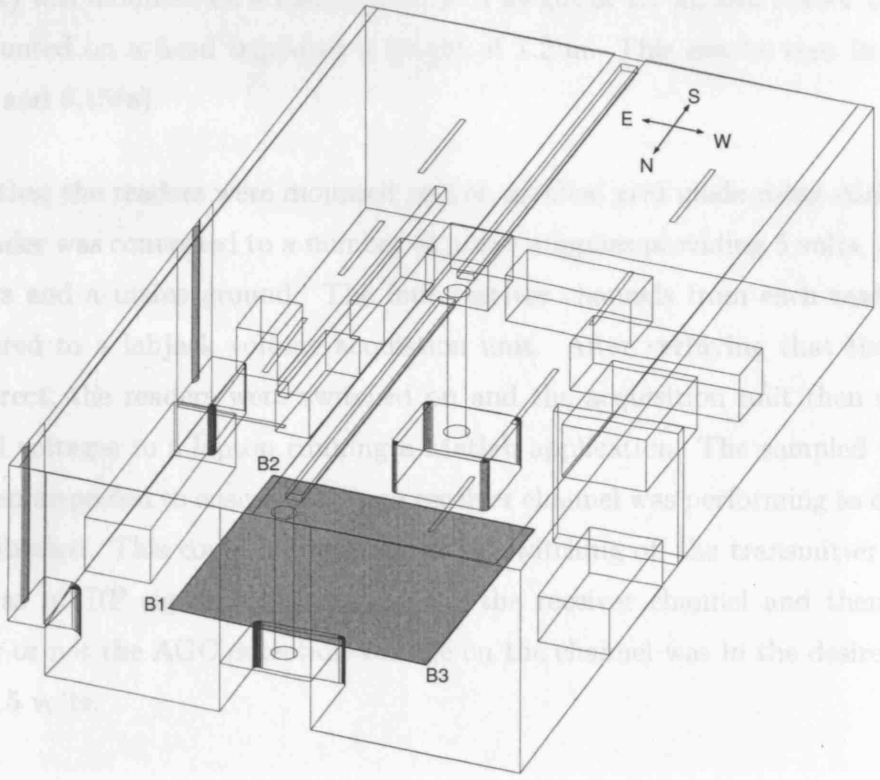
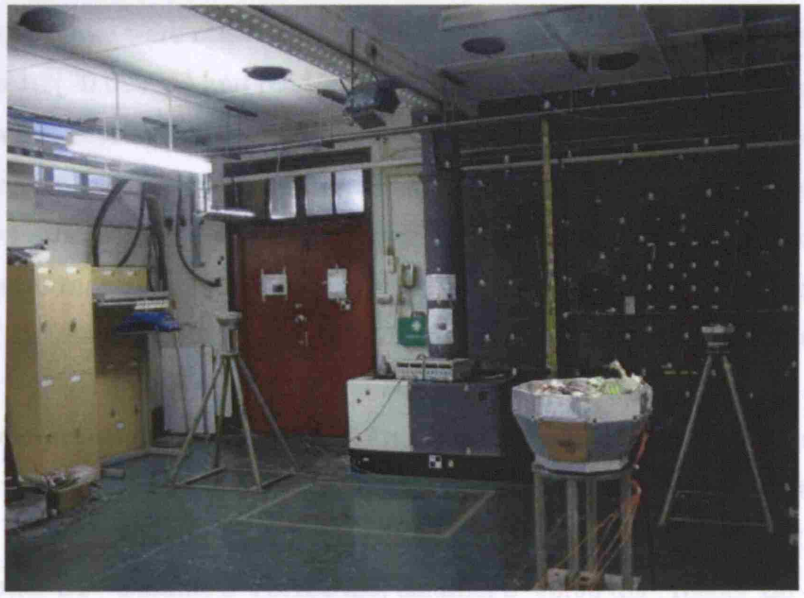


Figure 6.14: An isometric view of the location used for the trials in site B

An area between the two columns was selected for the experiments because it offered an open empty space unlike the remainder of room which was occupied by equipment and it also allowed the readers to be placed far enough apart in order to allow a meaningful number of location sample points. Now that the area was chosen a grid with a regular mesh was marked on the floor of the room using sticky tape. Having established the grid, the readers were mounted on the three vertice's of the grid forming a right angled triangle placement using fixed height tripods. Two of the tripods had the same height while the remaining tripod had a slightly lower height. As a consequence, reader B1 was mounted 20 cm lower than readers B2 and B3. Reader B1(0,0) was mounted on a fixed tripod at a height of 1 m, reader B2(0,4.2) was mounted on a fixed tripod at a height of 1.2 m, and reader B3(3.6,0) was mounted on a fixed tripod at a height of 1.2 m. This can be seen in Figures 6.15(b) and 6.15(a).

Now that the readers were mounted and an artifical grid made using sticky tape, each reader was connected to a number of power supplies providing 5 volts, 12 volts, -12 volts and a mains ground. The four receiver channels from each reader were hard wired to a labjack voltage acquistion unit. After verifiying that the wiring was correct, the readers were switched on and the acquisition unit then sent the sampled voltages to a laptop running a Matlab application. The sampled voltages were then inspected to ensure that each receiver channel was performing as designed and calibrated. This could be determined by switching off the transmitter so that there was no RF signal being received by the receiver channel and then noting whether or not the AGC detection voltage on the channel was in the desired range 2.2 to 2.5 volts.

(a) The far end of room away from the measurement grid

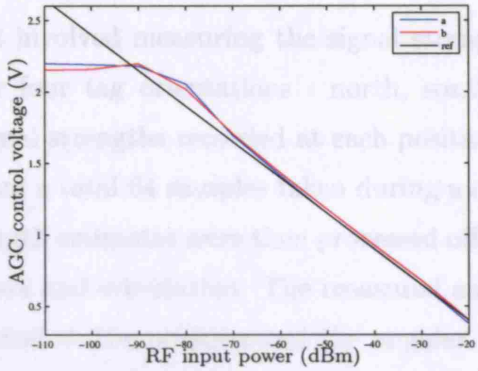


(b) The three readers and their respective mountings

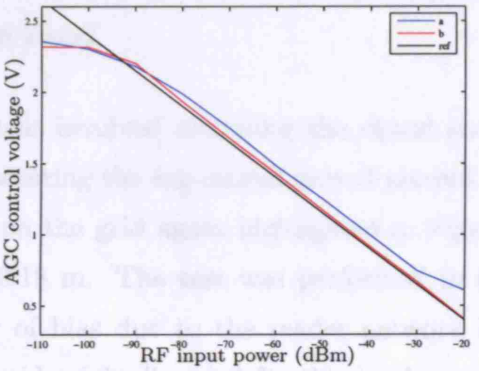
Figure 6.15: An indication of the experimental conditions within site B

6.2.1 Description of the measurement procedure

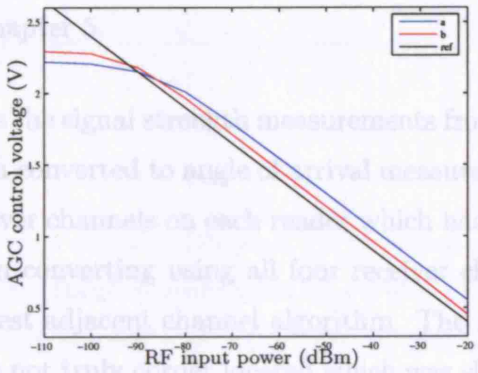
Before the sampled voltages representing the signal strengths could be used reliably in the experiment any ground offsets between the reader grounds and the labjack ground were measured so that they could be removed as required when the samples were processed offline. Since all the readers, the power supplies, and the labjack were connected to the mains ground at various points in the room it was a simple case of recording the ground offsets between the labjack ground and the ground returns from one of the receivers within each of the readers. The offsets to be measured were below the threshold of the acquisition unit and were therefore measured using a handheld AVO. Because the ground return current from each AGC board varies depending on how much amplification is occuring within the AGC loop, the ground offset was measured both when the transmitter was off and on. The offsets were all found to be less than 100 mV. Indeed it was found that there was only a small variation of ground offset for each reader when the transmitter was turned on and off. Average values were taken which resulted in ground offsets for reader B1(-0.096 V), reader B2(-0.029 V) and reader B3(-0.024 V). This procedure ensured accurate voltages were recorded. While the absolute voltage values are not necessary since relative voltage measurements were only required to derive the angle of arrival at each reader, the absolute voltage values could potentially be used to improve the location performance. Having accounted for grounding offsets, each reader were also calibrated to quantify any variation from the designed AGC characteristic of slope 40 dB per volt and 0.65 V @ -30 dBm input power. The results of the calibration for the receivers used in each reader are shown in Figure 6.16. Both readers B2 and B3 contain receivers that have characteristics which are linear in the desired operating region, -80 dBm to -30 dBm input power and contain fixed AGC voltage offsets, reader B2 - channel a +3.152 dB, channel b +0.96 dB ; reader B3 - channel a 5.8 dB, channel b +2.36 dB. Reader B1 contains receivers which match the desired characteristic fairly well from -65 dBm to -30 dBm and contain the fixed offsets, B1 - a zero offset, b +1.2 dB. At lower input powers of -80 dBm the voltage offsets from the desired characteristic become as much as 6 dBm which suggests a deviation from the linear characteristics of the AGC as a result of loop instability.



(a) Calibration of reader B1



(b) Calibration of reader B2



(c) Calibration of reader B3

Figure 6.16: Comparison of the calibration of the three pairs of receivers used in the experiments in site B with respect to the reference response with the (signal strength (nominal)= -(AGC Control Voltage - 0.65)*40 -30 dBm)

6.2.2 Outline of the experimental tests to be conducted in site B

The first experiment involved measuring the signal strength at a number of positions on the grid for four tag orientations - north, south, east and west. For experiment one the signal strengths recorded at each position are an average signal strengths measured from a total 64 samples taken during a one second sampling interval. The signal strength estimates were then processed offline to obtain measured bearings at each position and orientation. The measured angle was compared with the actual angle of arrival to the position and the angular bearing error recorded. The grid shown in Figure 6.17 was 4.2 m wide by 3.6 m long and had a mesh size of 60 cm. This allowed a total of 53 location sampling points on which the person wearing the transmitter stood.

The second experiment involved recording the signal strength on each receiver channel as the person wearing the tag moved moved around a diamond path placed approximately central on the grid again highlighted in Figure 6.17. The total distance of the path is 10.18 m. The test was performed in two directions so as to remove any possibility of bias due to the reader network being located on three vertices of the square grid originally used for the north, south, east and west tests in experiment one. The data samples were then processed using Algorithms I, II, and III discussed in Chapter 5.

For both experiments the signal strength measurements from the receiver channels of the readers were then converted to angle of arrival measurements. The conversion used only the two receiver channels on each reader which had direct coverage of the transmitter rather than converting using all four receiver channels on each reader by applying the strongest adjacent channel algorithm. The reason for doing this is because the readers are not truly corner located which was shown to perform better in the previous experiments in site A and therefore including those channels which do not have a direct of sight could cause quadrant errors because of strong reflection off walls and the metallic structures nearby.

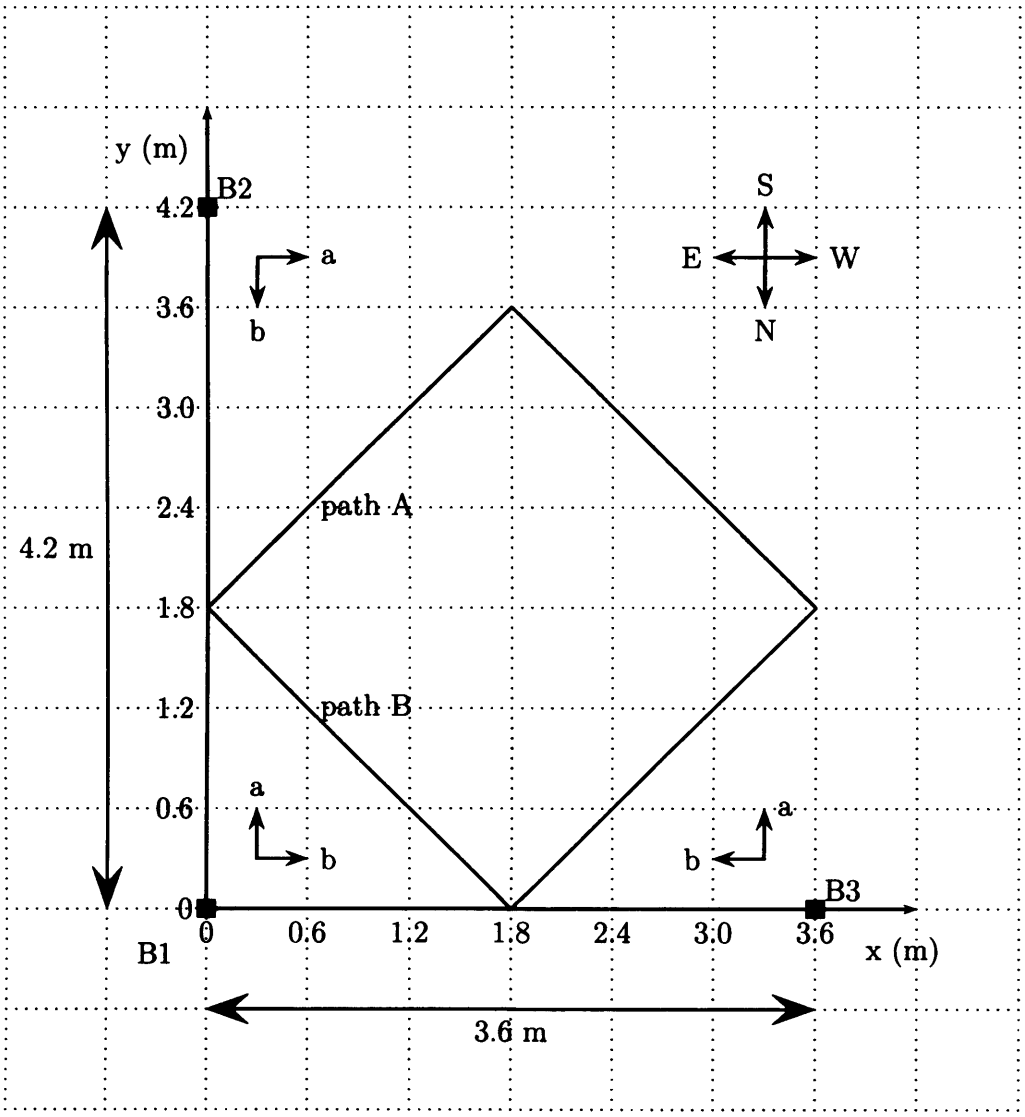


Figure 6.17: A plan view of measurement grid (1) experiment one has 8 by 7 sample points minus the three reader positions while (2) experiment two uses a diamond path starting at (3.6,1.8) with path A (clockwise movement) and path B (anti-clockwise movement)

6.2.3 Experiment one

The signal strengths at each transmitter position on the grid were processed and the average angular bearing errors calculated using the voltage ratios $20\log_{10}(\tan\phi)$ and $20\log_{10}(\tan^2\phi)$. The results are shown in Table 6.5 with the transmitter azimuthal pointing orientations that generated the lowest mean angular error, μ and standard deviation, σ, highlighted in yellow for each reader. The lowest angular error present was from reader B3 with a mean angular error, μ, of $-1.56°$ and a standard deviation, σ, of $12.11°$ when the tag was pointing north whilst the largest angular error present was from reader B2 with a mean angular error, μ, of $-18.29°$ and a standard deviation, σ, of $23.29°$ when the tag was pointing west. In addition the table also shows a comparison between the angular bearing errors calculated with and without the measured offsets from the designed linear in dB response that were found in an earlier calibration for each receiver. The first observation that can be made from the results is that when we compare the angular bearing errors calculated from receivers with a reference response and when calculated from receivers with the calibrated response we find that there is no dramatic improvement in the angular bearing accuracy from the calibration of the receiver channels. In fact the calibration has caused little change to the angular bearing errors found with reader B1 but surprisingly has added a small amount of bias to the mean angular errors from readers B2 and B3. It can also be observed that when we compare the angular bearing errors for both voltage ratios we find that both functions provide very similar angular bearing error with neither function providing distinctly better accuracy than the other function.

Having established the effect of calibration and the voltage ratio functions on the angular bearing error, we now look at the performance of the readers with respect to the pointing angle. Beginning with reader B1 we see that the angular bearing errors are lower in the east and north tag pointing angles. The east tag pointing angle provides a lower bearing error than the north pointing angle due to the metallic pillar and other metallic artifacts located on the northern perimeter of the room. Specifically the eastern pointing angle contains mean angular bearing error, μ, no

greater than 7.05° and standard deviation, σ no greater than 14.49°. This positive bias in the mean could be attributed to the level of reflection from the northern face of the room being greater than the level of reflection from the eastern face of the room. As a result, the channel receiving the lower level of reflected signal is more likely to be stronger than the channel receiving the higher level of reflected signal due to the process of destructive interference. Looking now at the other two readers we see that reader B2 shows the lowest angular bearing error when the tag is in the south pointing angle. Specifically the southern pointing angle contains mean angular bearing error, μ, no greater than $-8.96°$ and standard deviation, σ no greater than 18.90°. Reader B3 shows the lowest angular bearing error when tag is in the north pointing angle. Specifically the northern pointing angle contains mean angular bearing error, μ, no greater than $-9.20°$ and standard deviation, σ no greater than 16.31°.

Having analysed the data it is clear that a certain level of bias is prevalent in all three readers but this is unavoidable since the readers are not truly corner located and there is a significant number of unmoveable reflecting metallic artifacts along the periphery of the measurement grid. Whenever the transmitter is in an unfavourable pointing angle the mean and standard deviation is much higher and leads to larger location errors. The size of this location error could be estimated in a similar way to the Debrecen measurement by multiplying the operating range by the tan of the rms bearing error from a single reader. While this affords a quick estimate of location error from a single reader it does not take into account the fact that the bearing errors from the readers may not be multiplicative.

Table 6.5: Experiment one - mean and standard deviation of the angular error in (°) for reader locations B1, B2, and B3 under four different transmitter pointing angles (key: ref = nominal uncalibrated measurements, cal = calibrated measurements)

Tag orientation	$\tan\phi$ ref		$\tan^2\phi$ ref		$\tan\phi$ cal		$\tan^2\phi$ cal	
	μ	σ	μ	σ	μ	σ	μ	σ
Reader B1								
North	14.39	16.40	8.62	17.80	15.53	16.01	9.30	17.71
South	10.08	19.95	4.76	20.96	10.47	19.87	4.83	21.27
East	5.61	10.58	2.27	14.49	7.05	10.81	3.09	14.96
WestB1	-6.81	25.09	-5.21	24.44	-4.80	24.70	-4.03	24.38
Reader B2								
North	5.77	26.35	2.48	25.14	-0.78	26.62	-1.16	25.63
South	-2.23	16.40	-2.06	18.24	-8.96	16.60	-6.48	18.90
East	-8.59	20.06	-6.07	20.23	-14.65	20.49	-10.09	20.38
West	-12.88	23.56	-8.80	22.85	-18.29	23.29	-12.89	22.80
Reader B3								
North	-1.56	12.11	0.35	15.88	-9.20	13.26	-4.73	16.31
South	-4.61	26.37	-1.46	24.55	-12.55	26.40	-6.45	24.73
East	1.35	29.38	2.46	26.76	-7.57	29.40	-2.94	26.80
West	0.21	22.79	1.81	22.07	-7.85	22.47	-3.30	22.02

The measured bearings from the uncalibrated ($\tan\phi$ ref.) function were then combined using the least squares algorithm discussed in Chapter 5. The measured location estimates from the least squares algorithm were then compared wih the actual location positions. In order to quantify any bias in the location error the the individual location errors are averaged in the x direction, Δ_x and the y direction, Δ_y, respecively and then combined to provide a signed location error. To get an indication of the overall average location error that we can anticipate from the network we then calculate the absolute location error. The location errors for each tag orientation are shown in Table 6.6 and also the errors in Δ_x and Δ_y are also illustrated in Figures 6.19 and 6.20. The tag orientation with the lowest absolute error is the east pointing angle with a mean, μ, of 1.16 m and standard deviation, σ, of 0.55. This tag orientation also has the lowest errors in Δ_x and Δ_y which are also illustrated in Figures 6.19(c) and 6.20(c). The angular errors are larger for the remaining three tag orientations with the west pointing angle having the largest combined error in Δ_x with a mean, μ, of 0.47 m and standard deviation, σ, of 1.03 m whilst the north pointing having the largest combined error in Δ_y with a mean, μ, of 0.49 m and standard deviation, σ, of 0.93 m. The cdfs of the absolute location error for the four tag orientations are plotted in cumulative density function of the absolute location errors as shown in Figure 6.18. The mean absolute location error for each tag orientation each correspond with the location error at the 50th percentile.

Table 6.6: Experiment one - average location error in x and y for the four tag orientations for (1) Δ_x (2) Δ_y, (3) signed error $\Delta_x + \mathrm{j}\ \Delta_y$, (4) absolute error $\sqrt{\Delta_x^2 + \Delta_y^2}$, using the measured bearings from readers B1, B2 and B3 which were computed using the uncalibrated logarithmic voltage ratio ($\tan\phi$)

tag orientation	Δ_x		Δ_y		$\Delta_x + \mathrm{j}\ \Delta_y$		$\sqrt{\Delta_x^2 + \Delta_y^2}$	
	μ	σ	μ	σ	μ	σ	μ	σ
North	-0.37	0.88	0.49	0.93	0.61	1.28	1.20	0.76
South	-0.21	0.85	0.52	1.05	0.56	1.35	1.29	0.60
East	-0.18	0.86	0.35	0.87	0.40	1.23	1.16	0.55
West	0.47	1.03	0.01	1.08	0.47	1.49	1.40	0.67

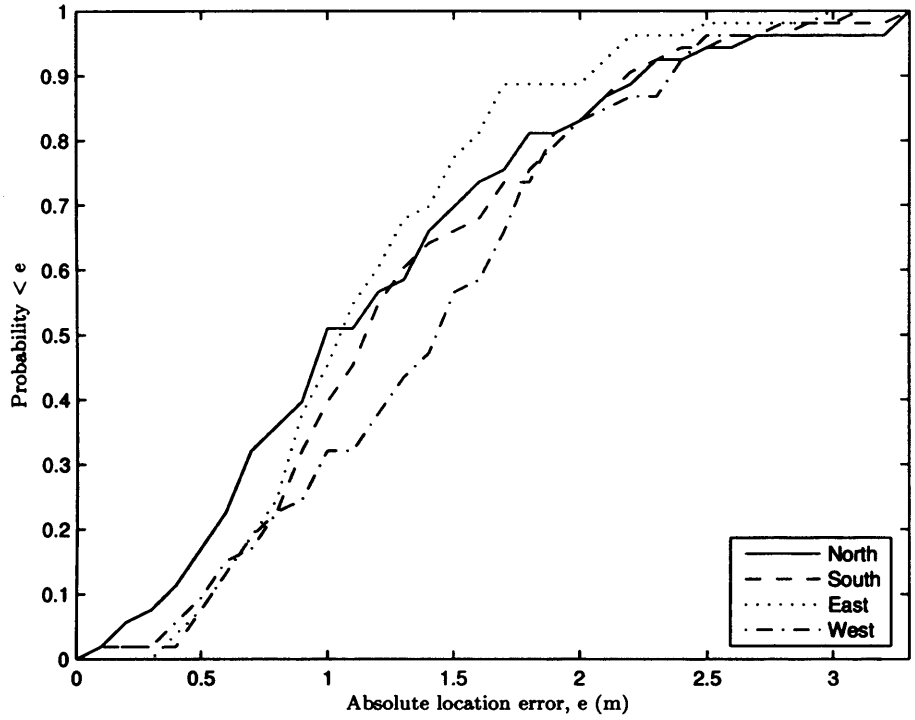
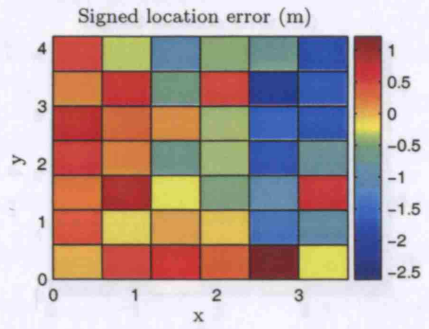
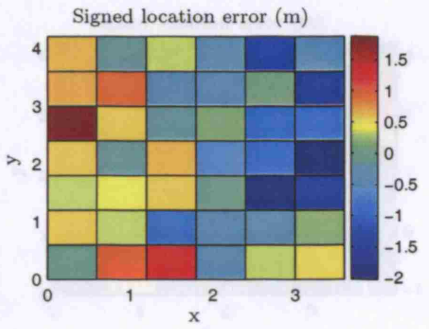


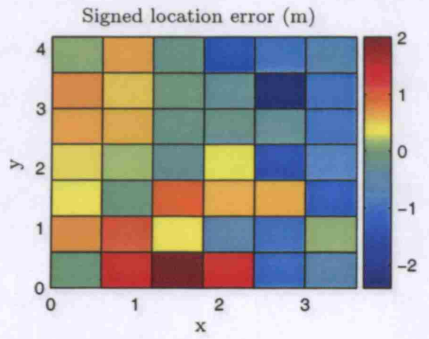
Figure 6.18: Cumulative density functions of the location errors for the four orthogonal tag orientations using the measured bearings from readers B1, B2 and B3 which were computed using the uncalibrated logarithmic voltage ratio ($\tan\phi$)



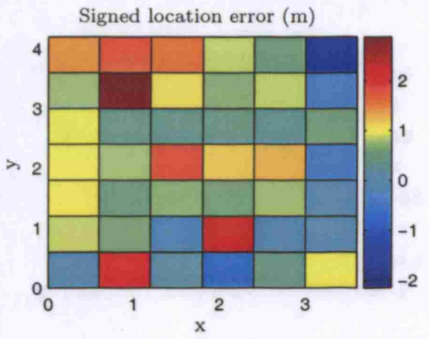
(a) Δ_x errors with the tag pointing (↓)



(b) Δ_x errors with the tag pointing (↑)

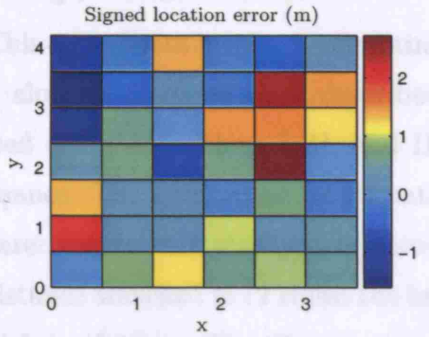


(c) Δ_x errors with the tag pointing (←)

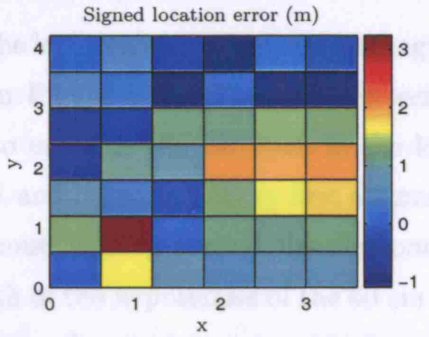


(d) Δ_x errors with the tag pointing (→)

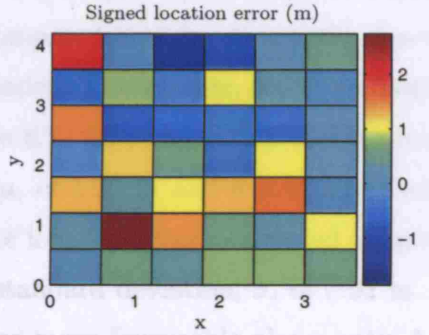
Figure 6.19: The location error in x, Δ_x, for the four tag pointing orientations using the measured bearings from readers B1, B2 and B3 which were computed using an uncalibrated logarithmic voltage ratio ($\tan\phi$)



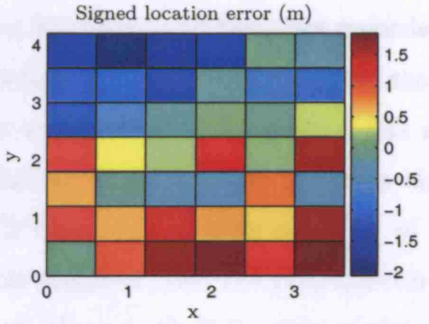
(a) Δ_y errors with the tag pointing (↓)



(b) Δ_y errors with the tag pointing (↑)



(c) Δ_y errors with the tag pointing (←)



(d) Δ_y errors with the tag pointing (→)

Figure 6.20: The location error in y, Δ_y, for the four tag pointing orientations using the measured bearings from readers B1, B2 and B3 which were computed using an uncalibrated logarithmic voltage ratio ($\tan\phi$)

6.2.4 Experiment two

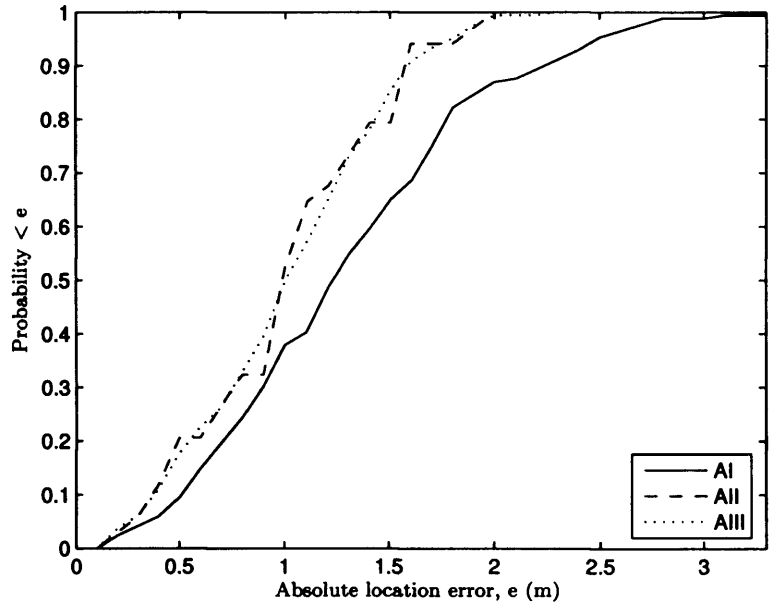
This experiment involved estimating the location error whenever the tag wearer walked along a diamond path described in Figure 6.17. The samples were then processed using Algorithms I, II, and III to examine the variation in the location performance. The location errors for path A and B are derived by first assuming the tag wearer moves at a steady rate throughout walking around the diamond. The total distance travelled is 12 times the length of the hypotenuse of the 60 cm square tile which is 10.18 m. The distance between each sample point is 10.18 m divided by the 170 samples which is approximately 6 cm. An array of location samples were then compared with location estimates using the measured bearings recorded from each reader. The location errors for both directions around the diamond are shown in table 6.7. The absolute location error from algorithm I for path A has a mean value, μ, of 1.29 m and a standard deviation, σ, of 0.65 which is better than the absolute location error calculated for path B which has a mean value, μ, of 1.74 m and a standard deviation, σ, of 0.92 m. This suggests that the propagation conditions are more favourable along path A than along path B for this reader layout. The use of averaging using algorithms II and II has improved the absolute location error for both paths of movement. This can also be seen in the cdfs of the absolute location error after applying algorithms II and III with a filter width, η, of 5 or 10. The performance of both algorithms are broadly similar with the best case location error being a mean of 0.92 m and standard deviation 0.34 m. The improvement in performance comes at the cost of less sample points along the path travelled by the tag wearer but this is tolerable as long as the tag repetition rate remains very large compared to the speed of movement.

The angular bearing errors found within site B at UCL are not too dissimilar to the corner mounted reader measurements from site A in Debrecen. Provided the propagating conditions are favourable, an accuracy of at least one metre can be expected at the typical operating range of 5 m. This can be seen from the average angular error for reader B3 as shown in Table 6.5. For a north tag pointing angle the reader has a mean angular error, μ, of $-1.56°$ and a standard deviation, σ of $12.11°$.

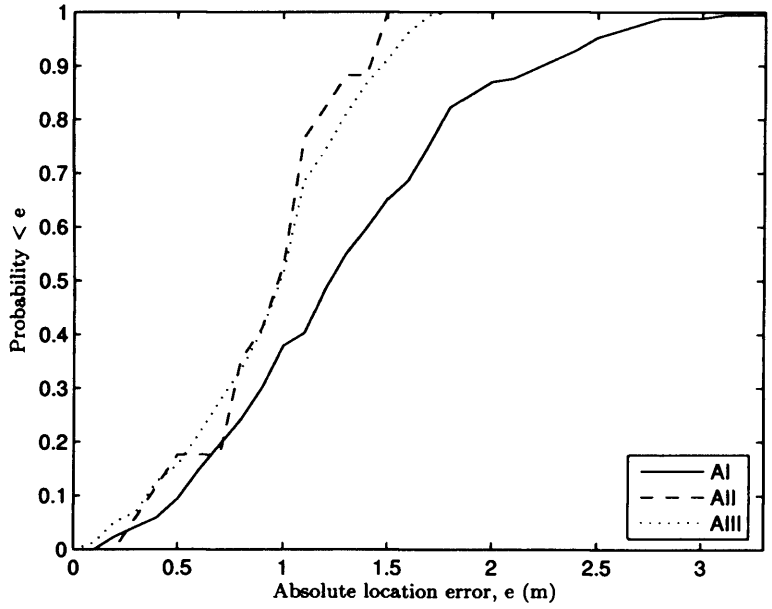
The propagation conditions can vary substantially depending on the tag location and pointing angle. Therefore the mean and standard deviation of the angular bearings from each reader can vary substantially as well and as a consequence so does the location accuracy. It is therefore more difficult to infer the likely location accuracy from a set of angular bearings taken separately. The measurements from experiment two establish the accuracy of the reader network when a tag wearer moves, with an approximate speed, along a path of known length. The location accuracy for path A was improved from a mean of 1.29 m and a standard deviation of 0.65 m using algorithm I to a location error as low as a mean of 0.92 m and a standard deviation of 0.34 m by filtering the measured angular bearings using algorithms II and III. The reduction in location error which can be achieved by averaging is dependent on the quality of the measurements.

Table 6.7: Experiment two - average location error in x and y for (1) Δ_x (2) Δ_y, (3) signed error $\Delta_x + \mathrm{j}\,\Delta_y$, (4) absolute error $\sqrt{\Delta_x^2 + \Delta_y^2}$, using the measured bearings from readers B1, B2 and B3 which were computed using an uncalibrated logarithmic voltage ratio ($\tan\phi$)

	Δ_x		Δ_y		$\Delta_x + \mathrm{j}\,\Delta_y$		$\sqrt{\Delta_x^2 + \Delta_y^2}$	
	μ	σ	μ	σ	μ	σ	μ	σ
Path A								
Algorithm I	0.17	0.86	0.45	1.06	0.48	1.36	1.29	0.65
Algorithm II $\eta = 5$	0.047	0.76	0.61	0.97	0.61	1.24	1.00	0.47
Algorithm III $\eta = 5$	0.17	0.60	0.47	0.82	0.49	1.01	1.01	0.46
Algorithm II $\eta = 10$	0.069	0.80	0.53	1.01	0.53	1.29	0.92	0.34
Algorithm III $\eta = 10$	0.17	0.54	0.50	0.79	0.53	0.95	0.94	0.40
Path B								
Algorithm I	0.17	0.86	0.45	1.70	0.48	1.91	1.74	0.92
Algorithm II $\eta = 5$	0.047	0.76	0.31	1.73	0.31	1.90	1.58	0.83
Algorithm III $\eta = 5$	0.17	0.60	0.46	1.61	0.49	1.72	1.59	0.83
Algorithm II $\eta = 10$	0.069	0.80	0.45	1.75	0.45	1.92	1.50	0.81
Algorithm III $\eta = 10$	0.18	0.54	0.48	1.59	0.51	1.68	1.56	0.81



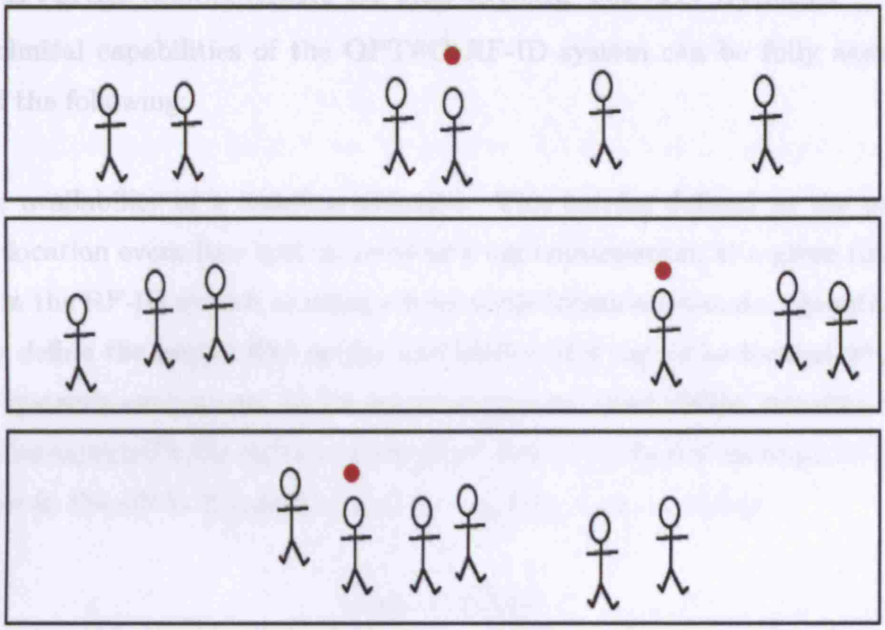
(a) Cdf comparison for a filter block size, η, of 5



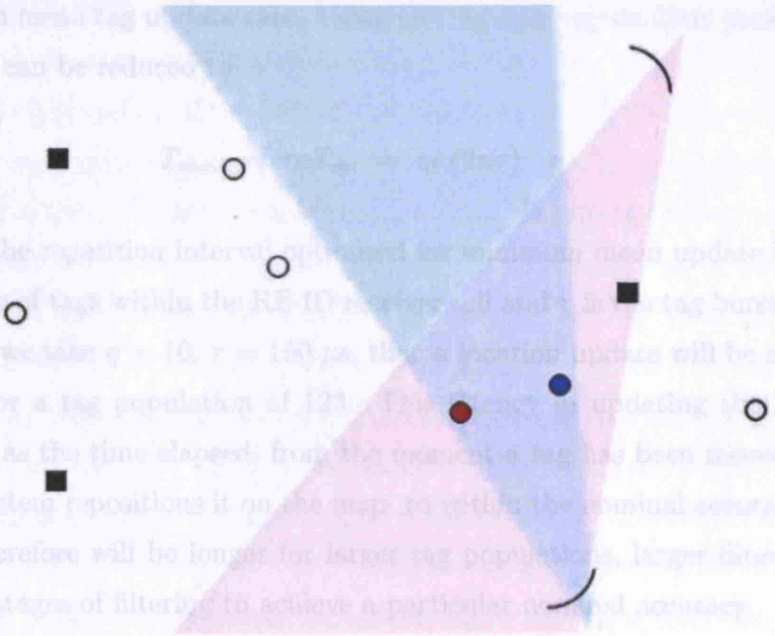
(b) Cdf comparison for a filter block size, η, of 10

Figure 6.21: Cumulative density functions of the location errors for the two tag paths taken by the tag wearer with the measured bearings from readers B1, B2 and B3 computed using an uncalibrated logarithmic voltage ratio ($\tan\phi$)

Overall, the AOA based technique used in the standalone RF-ID system gives adequate accuracy to at best a sector within the room. The accuracy of the measured angle of arrival suffers when the the LOS is not present over a prolonged number of sample bearings. Even though there are several ways to average the bearings from each reader, this makes it more difficult to adaptively track the transmitter. At a basic level the tracking algorithm could model the motion of the transmitter since the relative position between sensor and scene only changes incrementally and we know people do not teleport between places around a scene! However the tracking algorithm is wholly dependent on the quality of the measurements. If the statistics of the sensor measurement errors do not continuously have zero bias and fit a Gaussian error variance profile then it very hard to track a moving tag wearing individual with adequate latency and accuracy. Without an omni-directional tag field pattern the tracking of the RF-ID is likely to have variable accuracy and be far from satisfactory. A hybrid approach combining the independently developed active RF-ID system and panoramic video system has the potential to actively track individuals carrying a tag much more accurately than the standalone RF-ID system. The addition of tag IDs and location estimates from the tag readers removes the limitation imposed by a geometric search of the panoramic images from the camera system as illustrated in Figures 6.22(a) and 6.22(b). The person highlighted in the panaromic images can be tracked as long as the location estimates from the RF-ID system coincide with the region of uncertainty in the camera images.



(a) Camera user interface: multicamera in which we can identify the same person in each image



(b) Tag IDs and location estimates from the RF-ID system - readers denoted by black squares - could assist in the tracking of person highlighted in the panoramic images

Figure 6.22: An illustration of the potential for co-operation between the Optag active RF-ID system and the panoramic video imaging system

6.3 Overall assessment of the Optag RF-ID system

The technical capabilities of the OPTAG RF-ID system can be fully assessed by some of the following:

(1) The availability of a location estimate. This can be defined as the probability of a location event (the system receives a tag transmission, at a given time) that results in the RF-ID system availing a reasonable location estimate. Quantitatively, we may define the availability as the probability of a tag to be located within the RF-ID system's error range. In the prototye system, the location estimate for each tag can be expected to be no more than three metres of the real transmitter position as shown in the cdf in Figure 6.21 and be available with a latency

$$T_{avail} = \eta T_{mean}$$

where: η is the filter width applied to the angular bearings at each reader and T_{mean} is the minimum mean tag update rate. Using the tag clashing analysis presented in Chapter 3 this can be reduced to

$$T_{avail} = \eta e T_{opt} = \eta e (2n\tau)$$

where: T_{opt} is the repetition interval optimised for minimum mean update interval, n is the number of tags within the RF-ID receiver cell and τ is the tag burst length. For instance if we take $\eta = 10$, $\tau = 150$ μs, then a location update will be available every second for a tag population of 123. This latency in updating the location can be defined as the time elapsed, from the moment a tag has been moved to the moment the system repositions it on the map to within the nominal accuracy. The latency will therefore will be longer for larger tag populations, larger filter widths and additonal stages of filtering to achieve a particular nominal accuracy.

(2) The accuracy of the location estimate can be defined in probabilistic terms, as the error range, within which the tag is located with a 90% certainty. This also means that in any given time, 90% of the tags would be located within the nominal

error range. Using the site B measurements the analog location accuracy is found to be slightly less than 1.5 m as illustrated in the cdf from Figure 6.21.

(4) The tag battery discharge time is the time it takes a fully charged tag to discharge, under a normal usage scenario. This can be estimated for the Optag prototype tag in the following manner. First lets assume the maximum radiated power for each tag is 25 mW and second the tag electronics are deemed to be 10 % efficient therefore the total input power required is 250 mW. If we also assume the tag transmits a burst of 100 μs twice every second then the average power requirement of the battery can be given by

$$P_{dc\,av} = 2 \times 100 \times 250 = 50\ \mu\text{W}$$

Considering a voltage level supplied by the battery of 5 volts then the estimated average current consumption follows from the ohm's law P = VI and is equal to 10 μA. The capacity requirement can be subsequently calculated using the relationship, C = IT where T is the time expressed in hours, for a 24 hour period the battery capacity must be at least 0.24 mAhr. This capacity represents a very conservative estimate of the current consumption for the present iteration of the prototype tag since even though the duty cycle of the tag has been optimised for maximum throughput of tag IDs, the frequency source that consumes the majority of the current in the tag has been implemented using several discrete devices. Hence the tag is less efficient at converting DC power to an AC signal source in comparison to the ideal case of the devices on the tag being integrated into a single chip with few external components, a lower voltage requirement, say 3 volts. Since the tag is purely a demonstrator and we are examining 'the proof of principle' a high capacity lithium battery with capacity of 750 mAhr was used in the prototype tag which for convenience were operated in CW mode to allow uninterrupted testing of the angular accuracy of the reader arrays without repeated recharging. It should be noted the lower current consumption of an integrated tag could be supplied by either a button battery or a paper battery. Button batteries would be a cheap method of implementation since if they are purchased in large quantities they may cost only 1-3 cents each and

would be commercially viable as part of the finished product. On the other hand paper batteries would be an alternative more attractive and elegant solution since a slim battery is much more compatible with cheap disposable portable RF-ID tag technology.

The sole purpose of the Optag RF-ID system is to identify and locate the tag wearer. This is in contrast to a communication system that provides location as an added feature and so is heavily over engineered for these tasks. The Optag RF-ID system performs location estimation of the transmitter position by performing triangulation of the measured AOAs from a minimum of three readers. The AOAs are generated by differential RSSI between the strongest receiver channels in each reader. While this approach does not provide the location accuracy of the wide bandwidth TOA/TDOA system it performs much better than systems utilising current infrastructure such as WLAN. Table 6.8 shows a brief comparison of the Optag RF-ID system with existing RF based systems. A more thorough evaluation of existing RF local positioning systems can be found elsewhere [43]. The Optag system compares very well with the RSS based system called RADAR. Although we have highlighted installation costs such as cabling and indirect disruption to normal activity as a limitation in the Optag RF-ID system this would be present with other new systems as well. The system costs are not as severe as one would expect using direction finding arrays as the complexity and cost of the reader used in Optag can be much reduced without any tradeoff in accuracy as shown in the experimental results by adopting a two element receiver array rather a four element receiver array. These readers can be conveniently mounted on the perimeter of the RF-ID cell and a minimum of six receiver channels can then detect the location of the transmitter in two dimensions.

Table 6.8: Comparison of the Optag RF-ID system with existing RF based location technologies

Positioning system	Positioning method	Accuracy Precision	Deployment / Scale	Limitations
Cellular networks (GSM)	Base station triangulation	50 - 300 m	Density of cellular nodes	Only where cell coverage exists
GPS	RF time-of-flight lateration	1 - 5 m	24 satellites worldwide	Not indoors
RADAR Microsoft Research [10]	802.11 RSS location	3 - 4 m 50% probability	3+ bases	Training phase Method based on path loss
Optag RF-ID system	AOA triangulation using differential RSSI	1.5 m with 90% probability	3+ bases	Installation & system costs

Other important features of the Optag RF-ID system are its scalability and adaptability. The system may be installed in parts in a particular site rather than all at once. The covered area may initially be small (e.g. for testing purposes) and later on be extended. The system is also adaptable. It is straightforward to adapt the system to changes in the sites structural details. The sites topology, partition and object arrangement are, by nature, nonuniform, and nonstationary. These details must be also be carefully taken into account since they have a major impact on the systems accuracy. Moreover, to maintain accuracy over the entire site, the location system must be adapted from time to time (rearrange antennas, recalibrate parameters). This process can be designed to be simple with the Optag system.

7. SUMMARY & CONCLUSIONS

7.1 Summary

The first stage of the work involved the evaluation of the current state of the art regarding prior attempts at tagging and airport tracking of passsengers. The most significant of these attempts included the trial of short range RF-ID tags at Zurich airport, the bluetooth transceivers at Billund airport and the pilot trial of passenger tracking using GSM mobile phones by BA. These systems have shortcomings which are not apparent with the Optag RF-ID system. First of all the Optag RF-ID system has a read range that is much greater than the one metre or less typical of a passive tag and so the readers do not have to be sited very carefully to ensure a successfully read tag ID as would be the case with the system deployed in Zurich airport. Secondly, the RF-ID tag envisaged for Optag may be embedded in a boarding card at the very least and is therefore not as bulky as the bluetooth transceiver adopted in the trial at Billund airport. Thirdly, the reuse of existing telecommunications infrastructure in the airport, particularly the GSM network, as was attempted by BA is also an inadequate solution for passenger tracking and accelerated check-in because not all passengers wish to carry mobiles and those that do would have to opt-in to such service. The evalution of the current state of the art also included a review of existing RF-ID tagging technology. RF-ID tags utilise three primary frequency ranges: low 100-500 kHz, HF 10-15 MHz and UHF 850-950 MHz, and 2.4-5.8 GHz. The HF frequency band is popular and used widely in short range applications such as access control and cards. It is relative inexpensive but the use of a higher frequency would be more appropriate for the Optag RF-ID system since the tag antenna is of a smaller form than a counterpart antenna at a lower frequency. The rest of the background material discusses location techniques and a short study of channel propagation modelling which is used to assess the performance of the location system whenever multipath is present. There are four principal location techniques: RSS, AOA, TDOA, TOA. TDOA and TOA offer the best accuracy but are the most costly to implement. RSS is the cheapest to implement but its accu-

racy is extremely vunerable to propagation conditions affecting the path loss model. AOA chosen for the Optag project lies somewhere in between these two extremes as it offers better accuracy than RSS but is much more straightforward to implement than TDOA/TOA. The background material concludes with a discussion of channel propagation modelling which is used extensively in channel sounding. The models adopted in channel sounding generally are over complicated, in some cases a very large reflection order, greater than three, is generated prior to computing the received signal strength and as a consequence these models have very lengthy runtimes. A simpler and less time consuming approach is to assess the performance of the direction finding array using a finite number of reflection paths (reflection order no more three). The result is slightly less precise than the other models but is equally valid.

The next stage of the work considered the implementation of the active RF-ID system. First, a link budget was performed to determine the maximum spatial separation of the RF-ID readers and then the cell capacity was evaluated using clashing analysis of the communication protocol adopted within each RF-ID cell. It was found that to provide full coverage of a location the network of readers must be spatial separated by no more 20 m. Although it should be noted this represents a very conservative estimate of the spatial separation and much larger separations are expected when the propagation conditions are favourable. It was also found that the RF-ID readers can identify and locate as many 1600 RF-ID tags when using an asynchronous communication protocol. This represents a very large capacity and is more than enough to cover the reading rate and capacity required for an application, such as passenger tracking. Having established the read range and reader separation a novel RF-ID receiver architecture incorporating frequency diversity was considered. Not only does this architecture prevents in band frequency interference from other systems from disturbing the narrowband RF-ID system chosen for Optag it can provide a potentially lower tag cost because oscillators of low frequency stability are less expensive than stable fixed frequency PLL sources. The extensive investigation of the architecture included a detailed description of the syn-

chronistation technique and the hardware implementation including the design of the carrier frequency estimator using SAW (surface acoustic wave) dispersive delay lines. The performance of the carrier frequency estimator was also compared for two peak search algorithms. Algorithm I assesses the carrier frequency using the two highest spectral peaks in the BPSK modulated received signal whereas Algorithm II assesses the carrier frequency using the highest spectral peak and then the highest spectral peak in a band of twice the signal bandwidth centred on the first spectral peak chosen. The results indicate that Algorithm II has superior performance to Algorithm I at low signal to noise ratios, although the implementation of Algorithm I leads to slightly lower latency in the estimation.

The development of the RF-ID system to be used in the Optag trial was described next. First of all, the RF-ID tag was designed and developed. The design of the tag was complicated by the high division ratio caused unavoidably by a high carrier frequency and a low reference frequency used in the synthesiser. A considerable reduction in the reference sidebands was demonstrated analytically by the judicious use of higher order filtering. This was confirmed by measurement where the output from the prototype RF-ID tag was found to have a clean signal spectrum with no spurious signals close to the carrier and a phase noise of -120 dBcHz^{-1} @ 500 kHz offset. The synthesiser was then carefully modified to allow FSK modulation without disturbing the dynamics of the loop. It was essentially for the reliable operation of the RF-ID system that the peak frequency deviation and mean centre frequency of the transmitted data burst were matched to the FSK discriminator used in the RF-ID receiver. The design and development of the RF-ID tag was then followed by the design and development of the rather more challenging RF-ID receiver. Each receiver in the reader was designed with 80 dB of dynamic range from which the linear-in-dB RSSI was measured. The receiver is linear in dB in the range -30 dBm to -90 dBm. This means the reader can identify and locate the angular bearing to the RF-ID tag to distances as far as 400 m when a LOS exists! The measurement of the AOA can be performed accurately provided that any offsets from the nominal response are noted in the calibration of each receiver

channel. The direction finding properties of the four element antenna array were also investigated. The accuracy of the array was evaluated for elevation errors, articulation and additive Gaussian receiver noise. It was found that small changes in the elevation of the received signal when evaluated by the azimuthal power ratio the angular error is only significant outside the $\pm$ 10 dB range and is no more than 10° over a range $\pm$16°, in AOA. Ultimately the accuracy of the DF antenna system is realized by minimising the channel unbalance between receiver channels used in the adjacent strongest algorithm and maximising the signal to noise ratio of the signal strength estimates in each channel. The angular errors caused by channel unbalance may in large part be removed by calibration or reduced in size by ideally using an antenna element with better articulation. Having established the theoretical limits of the direction finding array an inset fed rectangular patch was then designed with a $\cos\phi$ field pattern. The measured co-polar H-plane logaritmic power ratio in dB for the isolated antenna element corresponded well with a $20\log_{10}(\tan\phi)$ function.

The underlying behaviour of the passive direction finding array using ray tracing techniques was then established. The performance of the array in an indoor environment was assessed numerically by applying modelling in progressing degrees of complexity starting with model A an optimistic amplitude-only two dimensional model using an omni-directional transmit antenna element field pattern. The model development proceeded to model B which is a more complex and realistic three dimensional model. Not only does it take account of a transmit antenna with a directional $\cos\phi$ field, it also allows for polarisation loss, height differences between the transmitter and receiver, receiver tilt and changes in electric field polarisation and amplitude on reflection off dielectric surfaces. The propagation conditions in an enclosed environment present a considerable challenge for the array since multipath signals are present and there is not always a line of sight to transmitter. The results from model A suggest that the angular error is independent of reader mounting position and the location error may have sub-metre resolution assuming a line-of-sight of signal exists between the transmitter and reader. Obviously, larger variations in the angular error are more probable for different reader mountings and placements

if this assumption does not hold as was shown in the results from model B. Those results also suggest that the angular bearing error is likely to be lower for a two antenna element corner mounted reader than a four antenna element central mounted reader. Having assessed a single reader for angular accuracy, four reader topologies were then evaluated for location accuracy using three averaging algorithms. It was found that location error post averaging is reduced by using reader topologies of higher reader order because these topologies are more likely to have at least two readers with a line of sight to the transmitter. Without a significant improvement in the propagation conditions and also the quality of the reader measurements, as was seen in the results from model B, it is not easy to mitigate factors such as elevation error, multipath and reader distance adjustment of the weighting matrix.

The remaining part of the work presented experimental results of the deployment of the RF-ID system in two locations. The first location was at site A in the departure area of Debrecen airport. The performance of the RF-ID system was assessed using a series of tests. The mean and standard deviation of the angular error were taken for both a central mounted and corner mounted reader. These were performed when there were no obstructions on the grid and the tag wearer was positioned on a regular grid for four tag orientations. The corner reader mounting showed lower angular error than the central reader mounting. In fact when the propagating conditions were favourable, the mean angular errors was nearly zero and standard variation of the angular errors was 16° which equates to 1.4 m of location error at the typical operating range of 5 m. At site A obstruction tests were also conducted. The first test involved assessing the affect of crowding using eight people surrounding the tag wearer. When compared with unobstructed errors there was a slight increase in the angular error. The second test involved the recording of a series of channel voltage samples on a single RF channel with the direct path from the transmitter to the reader obstructed by people in line. This test showed that obstruction by one or more persons causes a significant reduction in signal strength - around 10 dB, but still provides an adequate level of signal. Although the presence of more than one person in the direct path makes little difference

to the received signal. This indicates a single person causes a large attenuation of the direct path when using a transmission frequency as high as 5.8 GHz. The final test at site A was a range/path loss measurement over a 15 m range for four tag pointing orientations. The results show that using RSS as an indication of transmitter distance is very unreliable whenever the received signal does not contain a LOS signal. If the results are extended beyond the 15 m range tested here it is feasible that the RF-ID reader designed for Optag could identify and perform direction finding of tags located at distances as a high as 100 m. A log distance path loss model was then fitted to the signal strength measurements from the LOS antenna pointing orientation and was found to have a 18.20 dB rolloff with a fading factor of 2.05 dB. This results provides a very accurate model for the expected channel propagation condition with LOS for a room of this type at an operating frequency of 5.8 GHz. The second location used for the experimental results was at site B in a facility in UCL. As well as finding the mean and standard deviation of the angular errors from each reader, actual location errors were also measured using a network of three readers arranged on adjacent corners of a rectangle. The first experiment involved measuring the mean and standard deviation of the angular error taken for each reader. These were performed when there were no obstructions on the grid and the tag wearer was positioned on a regular grid for four tag orientations. Provided the propagating conditions were favourable, the mean was $-1.56°$ and standard variation was $12.11°$ which equates to 1.07 m of location error at the typical operating range of 5 m. The second experiment combined the angular bearings from the three readers to establish the actual location accuracy of the reader network when a tag wearer moved, with an approximate speed, along a path of known length. The location accuracy for path A was improved from a mean of 1.29 m and a standard deviation of 0.65 m using algorithm I to a location error as low as a mean of 0.92 m and a standard deviation of 0.34 m by filtering the measured angular bearings using bearing averaging algorithms II and III. The thesis then concluded with an brief assessment of the Optag RF-ID system taking into consideration not only capabilities such as location availability, location latency, location accuracy, and tag battery discharge time but also with some comments on how the system

compares with other RF location systems including scalability and adaptability.

7.2 Conclusions and future work

The RF-ID system which has been described in this thesis is a customised system which can identify and track people in an indoor environment, such as modern airports. The customised system has been created since existing systems only have some of the features required for this challenging applications and they are usually more complex than desired. The design of the system has been considered in detail here since it is often overlooked in the presentation of the other location finding systems and is therefore not factored into the overall adequacy of those systems. It is imperative that the system is sufficiently tailored to the requirements of the application and the overall performance assessment of the RF-ID system must not consist solely of the location accuracy!

The detail design of the system has shown that the RF-ID system developed here can accommodate many thousands of tags simultaneously but also with the capability of high location accuracy. Indeed the experimental results from the trials of the RF-ID system in site A and site B have shown that the AOA location techniques used in each reader can certainly locate the transmitter to at worst a small sector of the RF-ID cell provided adequate averaging of the received bearings is applied. The results have also shown that a two element corner mounted reader suffers less from the multipath propagation present in an indoor environment and are capabable of a location error of only 1 metre at the typical operating range of 5 m. The location accuracy of the RF-ID system used here can be rivalled and bettered by other more sophisticated wide band techniques of location but is unlikely they will have the simplicity and robustness of the system adopted in this work. In general, the location accuracy demanded should also take into account what information is realistically achieveable, and what information is actually needed in places such as airports. For shops and toilet spaces, only entries and exits might be monitored; for corridors, only the passage past wall-mounted readers.

Finally, the thesis has also explored innovative techniques to faciliate the possibility of cheap tags by referring the complexity of the system to the reader. The most significant of these innovations was the investigation of a novel RF-ID receiver architecture incorporating frequency diversity that is able to accommodate tags with poor frequency stability [5]. Although the technique was not demonstrated practically here it could represent an important feature of the RF-ID system if was to be commercialised at a later date.

At 5 GHz, attenuation in human tissue is very high. Hence an on-chest tag will not operate over more than about 180°. If vertically polarised, a simple dipole may provide that cover. If horizontally polarised, a horizontal beam width close to 180° might be achieved, in a flat tag, by a linear array of 3 or 4 dipoles, phased to the appropriate phase curvature. For greater compactness, one two-dipole array might be phased to form a beam slewed to the left and a similar one, underneath it, might be phased to form a beam slewed to the right. This antenna system would be on the outside of the ground-plane, and the tag electronics could be on the inside of the same ground-plane. This represents a neat solution but there is further scope for improvement with the tag antenna beamwidth. A wrist worn tag may readily provide a 360° but a system using this approach could encounter problems with changes in the field polarisation occurring sporadically with movement in the arm while the person is walking.

The production version of RF-ID tag could be a custom-designed chip. Hence the key criterion is not the cost of COTS components but the ease of IC processing (or possibly licensing). The cost target for the readers should include a highly-geared trade-off between extra reader cost and consequential tag savings. This can be achieved by adopting the novel wide band synchronising receiver. This receiver could be evaluated using a real prototype. The performance of the peak search algorithms could be then be assessed and compared with the analytic models and simulation results. The prototype could be developed using COTS components and future research could then focus on the feasibility and design of an integrated CFE.

Further work is required to compare this narrow-band AOA system and wide bandwidth (TOA, TDOA, UWB) location positioning systems side-by-side, to confirm where the advantages and disadvantages lie. For instance, the Optag RF-ID system could be compared with a TDOA system presently being explored in a project here at UCL. The project is called 'TINA' and reuses existing infrastructure using the principle of "RF over optical fibre". This comparison could be supplemented by improvements in the propagation models described in this thesis. The propagation models could be extended by using more realistic room structures. The improvement could include the ability to input building layouts, to handle multilayered building materials and to input surfaces which are not perpendicular or parallel to the coordinate system.

REFERENCES

[1] K. Finkenzeller. *RFID Handbook: Radio-Frequency Identification Fundamentals and Applications.* John Wiley and Sons, 2000.

[2] S. E. Sarma. Towards the five-cent tag. Technical report, MIT, http://www.autoidcenter.org/research/MITAUTOID, 2001.

[3] Annex - 'description of optag project', October 2003.

[4] P. V. Brennan. An rf-tag based video tracking and location system. *Optag system description*, March 2005.

[5] T.M. McCoy, P.V. Brennan, and R.J. Bullock. An rf-id receiver using saw filters for wideband synchronisation. *13th IEEE International Conference on Electronics, Circuits and Systems*, pages 1204–1207, 2006.

[6] R. G. Stansfield. Statistical theory of d.f. fixing. *J. IEE.*, 94, Part IIIa:762–770, 1947.

[7] Dragos Niculescu and Badri Nath. Vor base stations for indoor 802.11 positioning. *Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM*, pages 58 – 69, 2004.

[8] J. Hightower and G. Borriello. Location systems for ubiquitous computing. *Computer*, 34(8):57 – 66, 2001.

[9] M.Ni Lionel, Yunhao Liu, Yiu Cho Lau, and Abhishek P. Patil. Landmarc: Indoor location sensing using active rfid. *Wireless Networks*, 10(6):701 – 710, 2004.

[10] Paramvir Bahl and Venkata N. Padmanabhan. Radar: An in-building rf-based user location and tracking system. *Proceedings - IEEE INFOCOM*, 2:775 – 784, 2000.

[11] Georgia E. Athanasiadou and Andrew R. Nix. A novel 3-d indoor ray-tracing propagation model: the path generator and evaluation of narrow-band and wide-band predictions. *IEEE Transactions on Vehicular Technology*, 49(4):1152 – 1168, 2000.

[12] K.R. Schaubach and IV. Davis, N.J. Microcellular radio-channel propagation prediction. *IEEE Antennas Propag. Mag. (USA)*, 36(4):25 – 34, 1994.

[13] K. Rizk, J.-F. Wagen, and F. Gardiol. Two-dimensional ray-tracing modeling for propagation prediction in microcellular environments. *IEEE Trans. Veh. Technol. (USA)*, 46(2):508 – 18, 1997.

[14] M.C. Lawton and J.P. McGeehan. The application of a deterministic ray launching algorithm for the prediction of radio channel characteristics in small-cell environments. *IEEE Trans. Veh. Technol. (USA)*, 43(4):955 – 69, 1994.

[15] Simon Haykin. *Communication Systems.* 4th edition, 2001.

[16] A. D. Whalen. *Detection of Signals in Noise.* New York: Academic Press, 1971.

[17] David C. Rife and Robert R. Boorstyn. Single-tone parameter estimation from discrete-time observations. *IEEE Transactions on Information Theory*, IT-20(5):591 – 598, 1974.

[18] David C. Rife and Robert R. Boorstyn. Multiple tone parameter estimation from discrete-time observations. *Bell System Technical Journal*, 55(9):1389 – 1410, 1976.

[19] S. Holm. Optimum fft-based frequency acquisition with application to cospas-sarsat. *IEEE Trans. Aerosp. Electron. Syst. (USA)*, 29(2):464 – 75, 1993.

[20] Walid K.M. Ahmed and Peter J. McLane. Method for coarse frequency acquisition for nyquist filtered mpsk. *IEEE Transactions on Vehicular Technology*, 45(4):720 – 731, 1996.

[21] M.K. Nezami, R. Sudhakar, and H. Helmken. Dft-based frequency acquisition algorithm for large carrier offsets in mobile satellite receivers. *Electron. Lett. (UK)*, 37(6):386 – 7, 2001.

[22] P. V. Brennan. *Phase-Locked Loops: Principles and practice.* McGraw-Hill, 1996.

[23] D. P. Morgan. *Surface-Wave Devices for Signal Processing.* Elsevier, 1985.

[24] Yuxing Zhang, Junjie Gong, Xianwen Zhou, Yulan Wei, and J.D.N. Cheeke. High-resolution spectrometer systems based on saw chirp filters. *1997 IEEE Ultrasonics Symposium Proceedings. An International Symposium (Cat. No.97CH36118)*, vol.1:89 – 92, 1997.

[25] P. Hartogh and G.K. Hartmann. A high-resolution chirp transform spectrometer for microwave measurements. *Meas. Sci. Technol. (UK)*, 1(7):592 – 5, 1990.

[26] Mervyn A. Jack, Peter M. Grant, and Jeffrey H. Collins. Theory, design, and applications of surface acoustic wave fourier-transform processors. *Proceedings of the IEEE*, 68(4):450 – 468, 1980.

[27] S.J. Orfanidis. *Electromagnetic Waves and Antennas.* 2004.

[28] S. E. Lipsky. *Microwave passive direction finding.* SciTech Publishing, 2004.

[29] V. A. Borovikov and B. Ye. Kinber. *Geometrical theory of diffraction.* Institution of Electrical Engineers, 1994.

[30] H. E. Dempsey and C. M. Russel. *Ultra High Frequency Propagation.* John Wiley, 1953.

[31] C. A. Balanis. *Advanced Engineering Electromagnetics.* John Wiley and Sons, 1989.

[32] T. H. Laby G. W. C. Kaye. *Tables of Physical and Chemical Constants: And Some Mathematical Functions.* 14th edition, 1973.

[33] A. R. Von Hippel. *Dielectric materials and applications.* Wiley, 1954.

[34] O. Landron, M.J. Feuerstein, and T.S. Rappaport. In situ microwave reflection coefficient measurements for smooth and rough exterior wall surfaces. *1993 43rd IEEE Vehicular Technology Conference. Personal Communication - Freedom Through Wireless Technology (Cat. No.93CH3305-0)*, pages 77 – 80, 1993.

[35] *American Institute of Physics handbook.* McGraw-Hill, 3rd edition, 1972.

[36] R. A. Poisel. *Introduction to communication electronic warfare systems.* Artech House, 2002.

[37] D. J. Torrieri. Statistical theory of passive location systems. *IEEE Transactions on Aerospace and Electronic Systems*, AES-20(2):183 – 198, March 1984.

[38] David Y. Hsu. *Spatial error analysis: a unified application approach.* Wiley-IEEE Press, 1998.

[39] H. E. Daniels. The theory of position finding. *Royal Statisical society*, 1951.

[40] K.V. Nikolakopoulos, D. Anagnostou, C.G. Christodoulou, and M.T. Chryssomallis. Estimation of direction of arrival for coherent signals in wireless communication systems. *IEEE Antennas and Propagation Society, AP-S International Symposium (Digest)*, 1:419 – 422, 2004.

[41] Andrea Goldsmith. *Wireless communication.* John Wiley, 2005.

[42] T.S Rappaport. *Wireless communications principles and practises.* Prentice Hall, 2002.

[43] M.S. Wilcox. *Techniques for predicting the performance of Time-of-Flight based local positioning systems.* PhD thesis, University College London, September 2005.

APPENDIX A

Modeling and simlation of the carrier frequency estimator

A.1 RMS voltage of a Gaussian white noise process

The Gaussian white noise process has a probability density function that is uniformly distributed as illustrated in Figure A.1

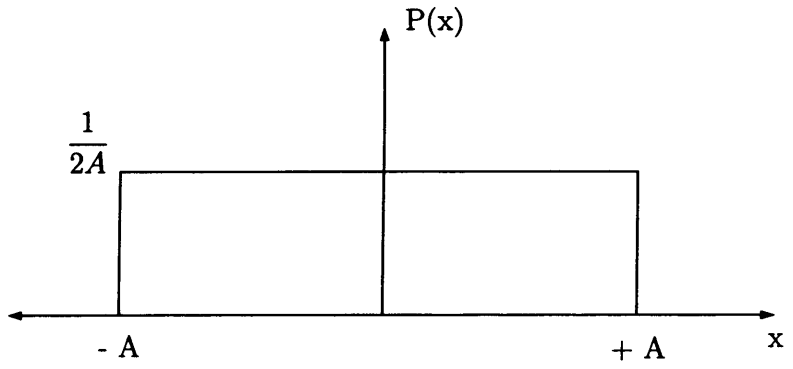


Figure A.1: Probability density function of a white Gaussian noise process

The variance of the noise process can be given by

$$\begin{aligned} \sigma^2 &= \int_{-A}^{+A} x^2 P(x) \\ &= \frac{1}{2A}\left[\frac{x^3}{3}\right]_{-A}^{A} \\ &= \frac{A^2}{3} \end{aligned} \tag{A.1}$$

The standard deviation, σ, of the noise process is therefore equal to

$$\sigma = \frac{A}{\sqrt{3}}$$

The scaling constant, A, is adjusted to set a desired noise spectral density, η_o, in the sampling bandwidth, $\frac{f_{sb}}{2}$.

A.2 Algorithm I analytic and simulation models

```
% ANALYTIC MODEL ALGORITHM I
ca=1; step =0.0005; Av=0:step:3;

for  s=[2^9]

Ndft=s; fres=fs/Ndft; N=(Ndft/2); R=1.25e6; B1=round(R/fres);
B2=round(2*R/fres);

measurement=0; SNR=LOGSPACE(-0.8,1,20); i=0;
measurement=measurement+1;

for i=1:length(SNR)

coverno = SNR(1,i)*R; no=ca/coverno; A= sqrt(ca)*(sqrt(2));
Pres=(no*fres); sigma=sqrt(Pres); b=(0.636*A); c=(0.212*A);

r1=(1-exp(-(Av.^2)/(2*(sigma^2)))).^((N)-4);
r2_a=(Av/(sigma^2)).*exp(-((Av.^2)+(c^2))/(2*(sigma^2)));
r2_b=besseli(0,(c.*Av/(sigma^2)))
r2= r2_a.*r2_b;
for j=1:length(r2);
area(j,1) = 0.5*step*((r2(1) +(r2(i)))+ (2*(sum(r2(2:(j-1))))));
end
r4= area'.^2;
r5_a=(Av/(sigma^2)).*exp(-((Av.^2)+(b^2))/(2*(sigma^2)));
r5_b=besseli(0,(b.*Av/(sigma^2)));
r5=r5_a.*r5_b;
r6=r1.*r4.*r5;
A1_s1=0.5*step;
A1_s2=((r6(1) +r6(length(r6)))+ (2*(sum(r6(2:length(r6)-1)))));
A1_s=A1_s1*A1_s2;
A1_fail = (1-((A1_s)^2));

end end
```

```
% SIMULATION ALGORITHM I
FFT=[ 2^9 ]; SNR=LOGSPACE(-0.8,1,20); Strial=3000; trial=0;
for trial=1:Strial; measurement =0;
for  s=1; measurement = measurement +1; p=length(SNR)
for j=1:p;
phase=0;fs = 80*1e6; Ts = 1/fs;fres(measurement,1)=fs/FFT(s);
R=2*625e3; fcs=20e6; R_N=R/fres(measurement,1);
fc_N(measurement,1) =(fcs/fres(measurement,1));
fc(measurement,1)=fc_N(measurement,1)*fres(measurement,1);
T = 1/R; N = FFT(s)/(fs*T); c=1; coverno =  SNR(1,j)*R;
no=c./coverno; d=2*([0;1;0;1;0;1;0;1;0;1;0;1;0;1;0;1]-0.5);
tarray = [0:Ts:N*T-Ts]';dmat = ones(T/Ts,1)*d';
dover1 = dmat(:); dover = dover1(1:FFT(s));
y = sqrt(2*c)*dover .* cos(2*pi*fc(measurement,1)*tarray + phase);
noise = sqrt(6*fs.*no)*(rand(length(y),1)-0.5);
y = y+ noise; Y = fft(y,FFT(s)); Y = (2*fft(y))/length(y);
Y(1) = 0; m=length(Y); fres=fs/m; f = fs*(0:m/2)/m;

[C,I] = max(abs(Y(1:m/2+1)));
A1(measurement,j,trial)=I; B1(measurement,j,trial)=C; Y(I)=0;
[C,I] = max(abs(Y(1:m/2+1)));
A2(measurement,j,trial)=I; B2(measurement,j,trial)=C; Y(I)=0;
if (A1(measurement,j,trial) ~= (fc_N(measurement,1)+1-(R_N/2)))
& (A1(measurement,j,trial) ~= (fc_N(measurement,1)+1+(R_N/2)))
    error1(measurement,j,trial) = 1;
elseif (A2(measurement,j,trial) ~= (fc_N(measurement,1)+1-(R_N/2)))
& (A2(measurement,j,trial) ~= (fc_N(measurement,1)+1+(R_N/2)))
    error1(measurement,j,trial) = 1;
elseif (A1(measurement,j,trial) == (fc_N(measurement,1)+1-(R_N/2)))
& (A2(measurement,j,trial) == (fc_N(measurement,1)+1+(R_N/2)) )
    error1(measurement,j,trial) = 0;
elseif (A2(measurement,j,trial) == (fc_N(measurement,1)+1-(R_N/2)))
& (A1(measurement,j,trial) == (fc_N(measurement,1)+1+(R_N/2)))
    error1(measurement,j,trial) = 0;
end end end end
Error1 = cumsum(error1,3);PE_sim1=(Error1(:,:,Strial)/Strial);
```

A.3 Algorithm II analytic and simulation models

```
% ANALYTIC MODEL ALGORITHM II
ca=1; step =0.0005; Av=0:step:3;
for  s=[2^9]; Ndft=s; fres=fs/Ndft; N=(Ndft/2);
R=1.25e6; B1=round(R/fres); B2=round(2*R/fres);
measurement=0; SNR=LOGSPACE(-0.8,3); i=0;
measurement=measurement+1;

for i=1:length(SNR) coverno = SNR(1,i)*R; no=ca/coverno;
A= sqrt(ca)*(sqrt(2)); Pres=(no*fres); sigma=sqrt(Pres);
b=(0.636*A); c=(0.212*A);
r1=(1-exp(-(Av.^2)/(2*(sigma^2)))).^(N-4);
r2=[(Av/(sigma^2)).*exp(-((Av.^2)+(c^2))/(2*(sigma^2))).*
besseli(0,(c.*Av/(sigma^2)))]; %
for j=1:length(r2);   .
area(j,1) = 0.5*step*((r2(1) +(r2(i)))+ (2*(sum(r2(2:(j-1))))));
end
r4= area'.^2;
r34=area';
r5=(Av/(sigma^2)).*exp(-((Av.^2)+(b^2))/(2*(sigma^2))).*
besseli(0,(b.*Av/(sigma^2)));
for j=1:length(r5);
area(j,1) = 0.5*step*((r5(1) +(r5(i)))+ (2*(sum(r5(2:(j-1))))));
end
r5_1=area';
q1_a=(r1.*r4.*r5.*r5_1);
q1_A=[0.5*step*((q1_a(1) +q1_a(length(q1_a)))+
(2*(sum(q1_a(2:length(q1_a)-1)))))];
r9=(1-exp(-(Av.^2)/(2*(sigma^2)))).^(B2-3);
r10=r9.*r5.*r34;
q2_A=[0.5*step*((r10(1) +r10(length(r10)))+
 (2*(sum(r10(2:length(r10)-1)))))];

AII_fail = 1 - (2*(q1_A*q2_A)); PE_an2(i,measurement)=AII_fail;
end end
```

```
% SIMULATION ALGORITHM II
FFT=[ 2^9 ]; SNR=LOGSPACE(-0.8,2); Strial=3000; trial=0;
for trial=1:Strial; measurement =0;
for  s=1; measurement = measurement +1; p=length(SNR)
for j=1:p;
phase=0;fs = 80*1e6; Ts = 1/fs;fres(measurement,1)=fs/FFT(s);
R=2*625e3; fcs=20e6; R_N=R/fres(measurement,1);
fc_N(measurement,1) =(fcs/fres(measurement,1));
fc(measurement,1)=fc_N(measurement,1)*fres(measurement,1);
T = 1/R; N = FFT(s)/(fs*T); c=1; coverno =  SNR(1,j)*R;
no=c./coverno; d=2*([0;1;0;1;0;1;0;1;0;1;0;1;0;1;0;1]-0.5);
tarray = [0:Ts:N*T-Ts]';dmat = ones(T/Ts,1)*d';
dover1 = dmat(:); dover = dover1(1:FFT(s));
y = sqrt(2*c)*dover .* cos(2*pi*fc(measurement,1)*tarray + phase);
noise = sqrt(6*fs.*no)*(rand(length(y),1)-0.5);
y = y+ noise; Y = fft(y,FFT(s)); Y = (2*fft(y))/length(y);
Y(1) = 0; m=length(Y); fres=fs/m; f = fs*(0:m/2)/m;
ttt=abs(Y(1:(m/2)+1));
[C,I] = max(ttt); AII_1(measurement,j,trial)=I;
BII_1(measurement,j,trial)=C; ttt(I)=0;
coor1=I+(R_N)+1; coor2=I-(R_N)-1; coor3=length(ttt);
ttt(coor1:coor3)=0; ttt(1:coor2)=0;
[C,I] = max(ttt); AII_2(measurement,j,trial)=I;
BII_2(measurement,j,trial)=C; ttt(I)=0;
if (AII_1(measurement,j,trial) ~= (fc_N(measurement,1)+1-(R_N/2)))
& (AII_1(measurement,j,trial) ~= (fc_N(measurement,1)+1+(R_N/2)))
error2(measurement,j,trial) = 1;
elseif abs(AII_1(measurement,j,trial) - AII_2(measurement,j,trial))
== R_N
error2(measurement,j,trial) = 0;
elseif (abs(AII_1(measurement,j,trial) - AII_2(measurement,j,trial))
~= R_N)
error2(measurement,j,trial) = 1; end end end end
Error2 = cumsum(error2,3);PE_sim2=(Error2(:,:,Strial)/Strial);
```

APPENDIX B
Integration of the AGC control voltage - RSSI

There are a number of alternate ways of dealing with the noise which may affect the channel power measurements. One way is to assume that the SNR can be increased by averaging the signals from a number of samples if the signals from successive samples are coherent and the noise is at least partially incoherent. The measured signal **x** can be represented as Equation B.1.

$$\mathbf{x} = \mathrm{x} + \xi \tag{B.1}$$

where: x is the true signal and ξ is the noise component with mean zero and standard deviation σ_n. The SNR is then given as Equation B.2.

$$\mathrm{SNR}_{\mathbf{y}} = \frac{|\mathrm{x}|}{\sigma_n} \tag{B.2}$$

If N measurements are acquired, the average signal is given as Equation B.3.

$$\mathbf{y} = \mathrm{x} + \frac{1}{N}\sum_{n=1} \xi_n \tag{B.3}$$

Assuming that the noise for each measurement is uncorrelated, then the SNR for the averaged scans is given as Equation B.4.

$$\mathrm{SNR}_{\mathbf{y}} = \frac{|\mathrm{x}|}{\sqrt{\mathrm{var}\left\{\frac{1}{N}\right\}\sum_{n=1}^{N}(\mathrm{x}+\xi_n)}} = \sqrt{N}\frac{|\mathrm{x}|}{\xi_n} = \mathrm{SNR}_{\mathbf{x}}\sqrt{N} \tag{B.4}$$

This shows that the signal to noise voltage ratio increases as the square root of the number of the averaged data samples.

APPENDIX C
Co-ordinate system and reflection theory

C.1 Antenna element - Euler Angles

A rectangular patch antenna with a beam pattern defined in the following co-ordinate system: The fringing fields are directed over the resonant ends of the patch. The electric field E-plane is directed along x and the magnetic H-plane is directed along the y axis. This is shown in Figure C.1.

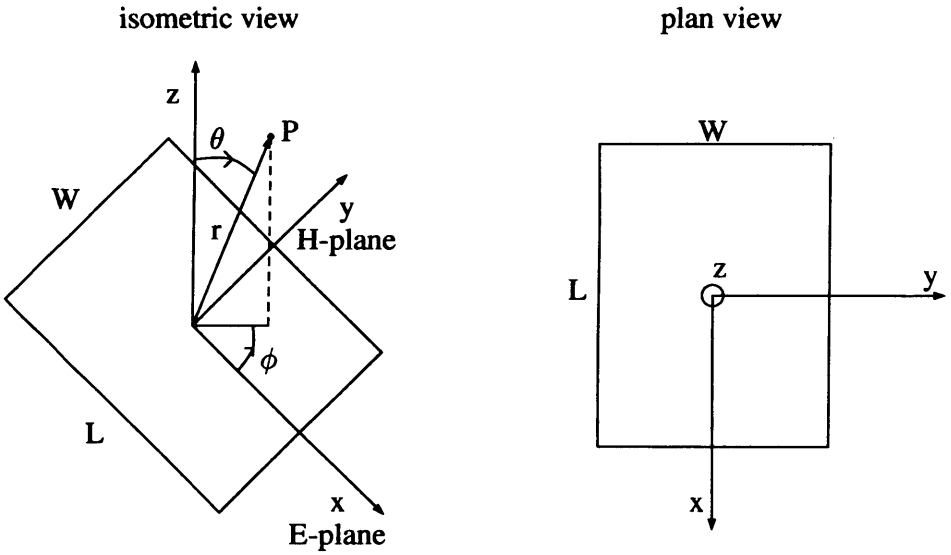


Figure C.1: The isometric and plan view of the co-ordinate system used for the microstrip patch antenna element

The normalised gain power pattern g is defined as a function r, ϕ, and θ. From the above co-ordinate system it can be seen that:

$$\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta \end{aligned} \tag{C.1}$$

and conversely,

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2} \\ \theta &= \cos^{-1}\left(\frac{z}{r}\right) \\ \phi &= \tan^{-1}\left(\frac{y}{x}\right) \end{aligned} \tag{C.2}$$

The problem is to plot the beam pattern for a vertically mounted patch using conventional azimuth (α) and elevation angles (ε), assuming a range r, as shown below in Figure C.2. This is a right handed co-ordinate system in x y and x.

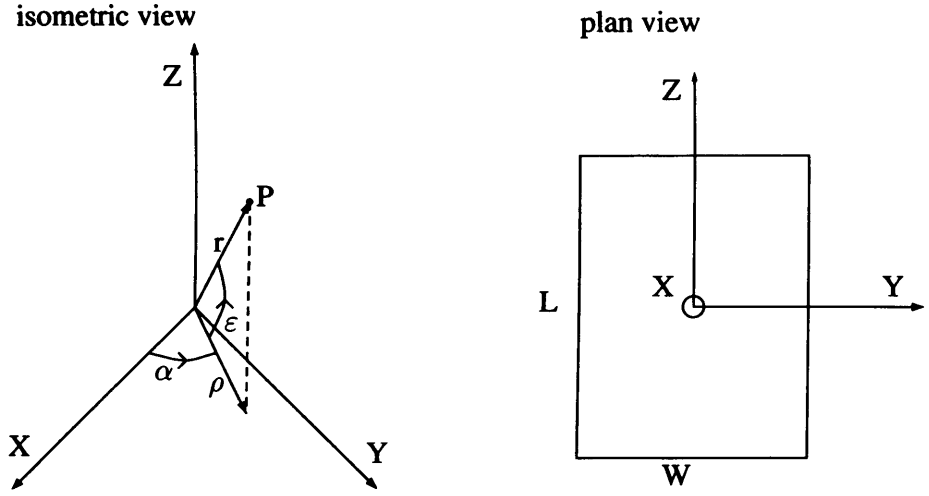


Figure C.2: The reference co-ordinate system for the angle of arrival of the received signals

$$\begin{aligned} X &= \rho\cos\alpha = r\cos\varepsilon\cos\alpha \\ Y &= \rho\sin\alpha = r\cos\varepsilon\sin\alpha \\ Z &= r\sin\varepsilon \end{aligned} \tag{C.3}$$

The conversion between the bearing reference co-ordinate system and the local patch co-ordinate system has the following steps:

1. Choose α and ε (fix r = 1).
2. Calculate P(X,Y,Z) in fixed co-ordinate system from (C.3).
3. Rotate P(X,Y,Z) to antenna body co-ordinate system P(x,y,z) (see below C.4).
4. Convert to antenna body spherical co-ordinates P(r,ϕ,θ) from (C.1).

5. Plot g(r,ϕ,θ) which is in fact g(r,α,ε).

The remaining problem is to express P(X,Y,Z), described in the fixed co-ordinate system as P(x,y,z) in a reference system aligned with the antenna which we may wish to align arbitrarily at a later date. To rotate from the fixed reference frame XYZ to a rotated frame xyz we may use the following 3 Euler angle rotations:

1. Rotate through Ψ_Z about the Z axis (for an aircraft this would be yaw).
2. Rotate through Θ_Y about the Y_{new} axis (pitch).
3. Rotate through Φ_X about the X_{new} axis (roll).

The rotations can be represented in matrix form as,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = [\Phi_X][\Theta_Y][\Psi_Z] \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \text{ where, } \Phi_X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\Phi_X & \sin\Phi_X \\ 0 & -\sin\Phi_X & \cos\Phi_X \end{bmatrix}$$

$$\Theta_Y = \begin{bmatrix} \cos\Theta_Y & 0 & -\sin\Theta_Y \\ 0 & 1 & 0 \\ \sin\Theta_Y & 0 & \cos\Theta_Y \end{bmatrix} \qquad \text{(C.4)}$$

$$\Psi_Z = \begin{bmatrix} \cos\Psi_Z & \sin\Psi_Z & 0 \\ -\sin\Psi_Z & \cos\Psi_Z & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Note the order of the matrices acting on the original untransformed co-ordinates. The transformations are not commutative. For the initial case of a vertically mounted patch with the H-plane (local y-axis) pointed at the target we need the following single rotation 90° about the Y axis shown in Equation C.5.

$$\Theta_Y = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{hence,} \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} x = -Z \\ y = Y \\ z = X \end{bmatrix} \qquad \text{(C.5)}$$

That this solution is correct can be seen by inspection of Figure C.3.

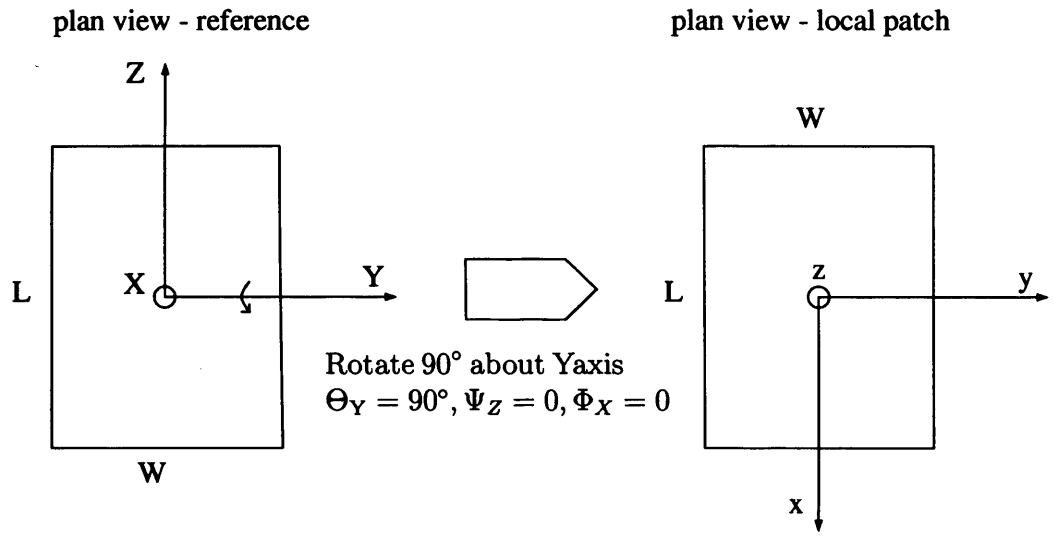


Figure C.3: Euler transformation between the reference co-ordinate system and the local patch co-ordinate system using a rotation about the Y-axis

However, if we extend the transformation to two antennas squinted in azimuth with the the same tilt angle the transformation equations to allow the examination of the power ratio are then not as trivial.

C.2 Image theory

The image position is formed from the intersection point of a vector normal to the surface **n** which passes through the transmitter position. The shortest distance from the point to plane the is shown in Equation C.6.

$$e = \frac{|(T - V) \bullet \mathbf{n}|}{|\mathbf{n}|} \tag{C.6}$$

where: T is the transmitter position with components (x_t, y_t, z_t) and V is a single vertex on the surface.

The transmitter image position $T'(x_t, y_t, z_t)$ can be found using Equation C.7.

$$\begin{aligned} x_t' &= x_t \pm 2e\left(\tfrac{a}{|a|}\right) \\ y_t' &= y_t \pm 2e\left(\tfrac{b}{|b|}\right) \\ z_t' &= z_t \pm 2e\left(\tfrac{c}{|c|}\right) \end{aligned} \tag{C.7}$$

where: the coefficients (a,b,c) describe the plane and the sign is determined by the position of the plane. For example in the room a wall at x = 0 m would have an image at a lower x value (minus sign), whereas a wall at x = 10 m would have an image at higher x value(plus sign). This principle can be evaluated numerically using the following function F_1 in Equation C.8.

$$F_1 = n \bullet T + d \tag{C.8}$$

where: d is the d is the perpendicular un-normalized distance of the plane to the origin, $-(\mathbf{n} \bullet V)$. If F_1 is less than zero sign then the sign is set plus in Equation C.7, however if F_1 is greater than zero then the sign is set minus in Equation C.7.

C.3 Plane line intersection

Since all surfaces have finite size it is necessary to determine whether a reflected transmitter image is visible to a receiver. This is done by checking whether the vector from the transmitter image to the base station passes through surface. From Figure (C.4) the position along a line L is calculated using the initial starting point P_0 added to a fraction k of the direction vector **u** which is calculated as the difference between two known points P_0 and P_1 along the line.

$$P = P_0 + k\mathbf{u} \tag{C.9}$$

The first test is the non-intersection test. If the dot product of the plane normal of the surface **n** and the ray direction vector **u** is equal to zero then there is no intersection of the plane and line. The plane and line can be then said to be parallel. The second test for disjointness or coincidence can be determined by testing any specific point which is contained in the plane π. That is whether it satisfies the implicit equation in (C.10)

$$\mathbf{n} \bullet (P - V_0) = 0 \tag{C.10}$$

If the line and plane are not parallel, then L and π intersect in a unique point P_{si}. At the intersect point, the vector $P_{si} - V_0 = \mathbf{w} + s\mathbf{u}$. This is equivalent to the dot

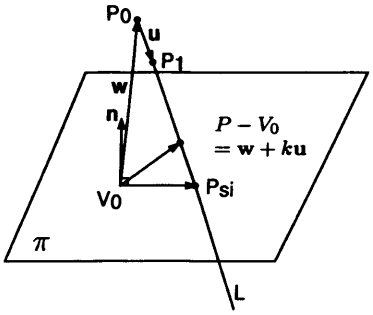


Figure C.4: Calculation of the plane line intercept for a ray through a surface

product condition: $\mathbf{n} \bullet (\mathbf{w} + k\mathbf{u}) = 0$. Solving we get (C.11):

$$k_{si} = \frac{\text{-}\mathbf{n} \bullet \mathbf{w}}{\mathbf{n} \bullet \mathbf{u}} = \frac{-(ax_o + by_o + cz_o + d)}{\mathbf{n} \bullet \mathbf{u}} \tag{C.11}$$

where x_o, y_o, z_o is the coordinates of position P_0 and a,b,c are the components of the unit normal vector, d is the perpendicular un-normalized distance of the plane to the origin. In the ray tracing model both points will be have been calculated and be positioned either side of the surface, for example P_0 is the transmitter image position, P_1 is the base station and **u** is a vector between the two points. Since the line is now a finite segment from P_0 to P_1 then one has just to check that $0 \leq k_{si} \leq 1$ to verify that there is an intersection between the line and the plane. The resulting intersection point P_{si} is then compared with the edges of the surface to check whether the point falls within the surface or not.